The PHILOSOPHICAL MAGAZINE

FIRST PUBLISHED IN 1798

A Journal of Theoretical Experimental and Applied Physics

Physicists, chemists and their biological colleagues have a fundamental common interest in the properties of molecules and molecular assemblies and the need has been growing for a journal which would bring together experimental and theoretical contributions to this field.

The physical aspects of molecules will be dealt with, in particular:

1. The structure of molecules as determined by physical methods and interpreted by quantum mechanics;

2. The electromagnetic properties of molecules and the processes of molecular excitation, ionization and dissociation;

3. The interaction between molecules and the equilibrium and transport properties of molecular assemblies.

The following contributions have been received for the first volume:

1. Nuclear Magnetic Resonance and Rotational Isomerism in Substituted Ethanes. By J. A. POPLE

2. Dielectric Properties of Iodine in Aromatic Solvents at 9000 Mc./sec. By G. W. NEDER-BRAGT and J. PELLE

3. On the Dipole Moments of Azines. By H. F. HAMEKA and A. M. LIQUORI

4. Electron–Electron Separation in Molecular Hydrogen. By M. P. BARNETT, F. W. BIRSS and C. A. COULSON

5. On Irreversible Processes in Quantum Mechanics. By I. PRIGOGINE and M. TODA

6. The Transport of Energy and Momentum in a Dense Fluid of Rough Spheres. By J. P. VALLEAU

7. Constant Pressure in Statistical Mechanics. By W. BYERS BROWN

8. One-dimensional Multicomponent Mixtures. By H. C. LONGUET-HIGGINS

Magnetohydrodynamic Shock Waves in the Solar Corona, with Applications to Bursts of Radio-frequency Radiation†

By K. C. Westfold

Department of Applied Mathematics, University of Sydney

Abstract

It is pointed out that the Type II burst velocities which are usually associated with the velocities of streams of ' auroral ' particles should rather be associated with the velocities of the shock fronts ahead of such streams. Likewise, the far greater Type III burst velocities can be associated with the shock velocities of corpuscular streams having similar velocities, if it is possible to raise the local value of the sound velocity to the magnitude of the Type III estimates. This can be effected if local magnetic fields of some hundreds of gauss are present in the solar corona, resulting in a predominant magnetohydrodynamic component in the sound velocity.

Starting from a set of transport equations previously investigated for an ionized gas, it is shown that in quasi-static electromagnetic fields the coronal medium satisfies an adiabatic equation of state and that the magnetic field lines are effectively ' frozen ' into the medium. The details of the transition across the shock front ahead of a stream in which all physical quantities depend only on the coordinate in the direction of flow is investigated for the case where the stream velocity is small compared with the sound velocity, and the results are compared with those for the hydrodynamic shock associated with a Type II burst.

The coronal conditions and the excitation of radiation near the shock front are briefly discussed.

§ 1. Introduction

Recent observations of the dynamic radio-frequency spectra of solar bursts (Wild and McCready 1950, Wild 1950 a, b) can be interpreted (Wild et al. 1953, Wild et al. 1954) in terms of disturbances which move outwards through the solar corona and excite the medium in their passage. The medium consequently emits radiation whose spectrum consists mainly of those components close to the electron plasma frequency that pertains at the level through which the disturbance is passing.

† Communicated by the Author.

On this interpretation, Type II bursts (Wild 1950 a), or outbursts, are due to disturbances which move outwards with velocities between 250 and 800 km/sec, and last for several minutes ; whereas the more common, usually bunched, Type III bursts (Wild 1950 b), or isolated bursts, are due to disturbances which move with velocities between 3×10^4 and 2×10^5 km/sec and last for several seconds. The actual movement of the sources of outbursts across the sun's disc has been established by the directional observations of Payne-Scott and Little (1952). By assuming plausible associated radial motions they were able to estimate the velocities as between 1000 and 3000 km/sec.

The present evidence (see Pawsey and Bracewell 1955, § 5.4) is consistent with the association of Type II bursts with corpuscular streams, such as are believed to accompany solar flares and to be responsible for non-recurrent magnetic storms and aurorae. The velocities of such streams range from about 800 to 1600 km/sec, which differ only by a factor of about 2 from the Type II velocities. The discrepancy can perhaps be ascribed to the possibly oblique passage of the disturbances through the surfaces of constant plasma frequency, and the possibility of local extensions of the corona above the regions of origin (Wild *et al*. 1953).

However, it appears on a closer scrutiny that the above considerations are too simple. It is known that the motion of a stream of particles, injected into a compressible medium, will be communicated to those particles of the medium that lie behind the front of a shock wave which moves with a velocity greater than both the stream velocity and the sound velocity in the undisturbed medium. Across the shock front there will be an abrupt transition, or jump, in the values of the physical quantities, e.g. velocity, density and pressure, that specify the state of the gas. It is likely that the disturbances at each level in the ionized solar atmosphere, occasioned by the passage of a shock front, will give rise to transient radiation, as in the cases considered by Jaeger and Westfold (1949), where it was shown that the prescription of initial values of the field vectors or the conduction current gives rise to an electromagnetic disturbance. The spectral components of the disturbance having frequencies below the electron plasma frequency of the medium are heavily attenuated while the other components are propagated with the appropriate phase velocity and attenuation, as given by Lorentz theory. It appeared that this transient radiation is also damped by collisions between the electrons and positive ions of the plasma. It may be noted here that Smerd (private communication) has shown that the general features of both Type II and Type III bursts can be accounted for by transient radiation emitted successively from each level.

It follows from these considerations that it is the accompanying shock velocity, rather than the corpuscular stream velocity, which should be associated with a Type II burst of radio-frequency radiation. Insertion of the value $\mu^2 = \frac{1}{4}$ (corresponding to 3 degrees of freedom for the

constituent gas molecules) in the formula (69.01) of Courant and Friedrichs (1948) gives for the shock velocity V ahead of a stream of particles moving with velocity v_1, into a medium where the sound velocity is a_0,

$$V = \tfrac{2}{3}v_1 + \sqrt{(\tfrac{4}{9}v_1{}^2 + a_0{}^2)}. \qquad \ldots \ldots \quad (1.1)$$

The sound velocity depends only on the temperature, for which the coronal value of 10^6 °K gives $a_0 = 170$ km/sec (see § 5). Thus, the shock velocity associated with a corpuscular stream moving through the corona at a speed of 1000 km/sec would be 1400 km/sec, a value still more discrepant from the Type II estimates, but within the range of outburst velocities estimated by Payne-Scott and Little. This new discrepancy can probably be accounted for as before, more easily if the molecular temperature is increased in the extended regions of the corona. On the other hand, if the shock velocity is taken as 500 km/sec the stream velocity would be only 330 km/sec, which corresponds to no other observed geophysical phenomenon.

The origin of Type III bursts has been regarded as more puzzling, since corpuscular streams having such high velocities could not be inferred from any associated solar or geophysical observations. However, Wild *et al.* (1954) have pointed out that the time delays between solar flares and the increases in terrestrial cosmic-ray intensity that occasionally accompany them, correspond to stream velocities of the same order as those associated with Type III bursts. Their observations show that clusters of these bursts may occur at the times of flares, sometimes alone and sometimes with a Type II outburst. The shock velocity corresponding to a stream moving with a velocity of 5×10^4 km/sec, which gives a delay of 50 minutes, in the corona at a temperature of 10^6 °K would be 7×10^4 km/sec, since, effectively, $V = \tfrac{4}{3}v_1$. This value is still within the range of Type III velocities.

On the other hand, once it is accepted that it is the shock velocity rather than the stream velocity that should be associated with both Type II and Type III bursts it becomes possible to account for the Type III velocities without having to postulate such large stream velocities. The velocity of the shock front ahead of a stream of particles, such as the auroral stream associated with Type II bursts, will rise to a Type III figure if the sound velocity is raised to about 10^4 or 10^5 km/sec. For the hydrodynamic shocks considered above, this would require temperatures of more than 10^{11} °K, which are very unlikely. However, a possible way out of the difficulty presents itself in the hypothesis (Chapman and Bartels 1940, p. 410) that the corpuscular streams that cause the aurorae and non-recurrent magnetic storms are emitted from M regions on the sun's surface, which occur on the sites of vanished sunspot groups and in association with the green coronal line.

There is thus a strong likelihood of the existence of magnetic fields of sunspot intensities in association with streams of auroral particles. In

such cases, as will be shown in § 3, the ' hydrostatic ' pressure will be augmented by the magnetic pressure and the internal molecular energy by the magnetic energy of the field, so that the sound velocity c_0 has a magnetohydrodynamic component. In fact c_0 is given by the formula of de Hoffmann and Teller (1950)

$$c_0{}^2 = a_0{}^2 + b_0{}^2, \qquad \cdots \cdots \qquad (1.2)$$

where b_0 is Alfvén's magnetohydrodynamic velocity in an incompressible medium. Since a magnetic field of 150 gauss in the lower corona gives $b_0 = 3 \times 10^4$ km/sec (see § 5), it appears that a plausible explanation of the large Type III velocities may lie in the hypothesis that this type of burst is associated with a magnetohydrodynamic shock wave, in contrast to the hydrodynamic shock waves associated with Type II bursts. Although it is generally supposed that such large fields do not reach into the corona, there are no measurements to support this view against any counter proposal. Indeed it will be shown in § 2 that the magnetic field lines move with the medium, so that the more intense fields near the surface can be convected into the corona.

Accordingly, the purpose of the present paper is to investigate the velocity of the shock front together with the jumps across the front of all the relevant physical quantities associated with a corpuscular stream. The results will be compared with those which derive from a hydrodynamic shock. After establishing some general magnetohydrodynamic relations of physical significance, the discussion is limited to the simplest case of rectilinear flow. Since, as will be shown in § 2, the electric conductivity of the corona is effectively infinite, this investigation will be found to include the detailed consideration of the case of a non-relativistic perpendicular longitudinal shock, which occurs in the general investigation of de Hoffmann and Teller (1950) and that of Lüst (1953).

§ 2. The Magnetohydrodynamic Equations for the Solar Corona

Approximations to the hydrodynamic equations, the equation of thermal energy, and the equation for the conduction current in a binary ionized gas have been previously investigated (Westfold 1953) using the iterative methods of Chapman and Enskog. In this paper only quasi-static electromagnetic fields will be considered, i.e. radiation will be neglected. It can be considered at a later stage, as a perturbation of the solution of a specific boundary-value problem in the present framework.

Since, in this case, conditions in the solar atmosphere permit the neglect of the free-charge density (Cowling 1953, p. 533), we may write

$$n_i e_i + ne = 0, \qquad \cdots \cdots \qquad (2.1)$$

where n_i, n are the number densities of the ion and electron components, and e_i, e the corresponding charges carried by the individual particles of the fully ionized medium, which is predominantly hydrogen. The

equations of continuity, mass motion and thermal energy then become

$$\frac{D\rho}{Dt} + \rho \frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{v} = 0, \qquad \ldots \ldots \ldots \quad (2.2)$$

$$\rho \frac{D\mathbf{v}}{Dt} = \mathbf{j} \wedge \mathbf{B} - \frac{\partial p}{\partial \mathbf{r}}, \qquad \ldots \ldots \ldots \quad (2.3)$$

$$\tfrac{1}{2} Nk(n+n_i) \frac{DT}{Dt} = -p \frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{v}, \qquad \ldots \ldots \quad (2.4)$$

where

$$p = (n+n_i)kT. \qquad \ldots \ldots \ldots \quad (2.5)$$

In these equations ρ is the mass density, $\mathbf{v}$ the mass velocity, $\mathbf{j}$ the conduction current, $\mathbf{B}$ the magnetic induction†, p the 'hydrostatic' pressure and T the kinetic temperature of the gas at the position $\mathbf{r}$ and time t ; k is Boltzmann's constant and N the number of degrees of freedom per particle, in this case 3 ; as usual, the differential operator $D/Dt = \partial/\partial t + \mathbf{v} \cdot \partial/\partial \mathbf{r}$. The displacement current may be neglected, so that Maxwell's equations give

$$\frac{\partial}{\partial \mathbf{r}} \wedge \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0, \qquad \ldots \ldots \ldots \quad (2.6)$$

$$\frac{\partial}{\partial \mathbf{r}} \wedge \mathbf{B} = \mu \mathbf{j} \qquad \ldots \ldots \ldots \quad (2.7)$$

and

$$\frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{E} = \frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{B} = 0, \qquad \ldots \ldots \ldots \quad (2.8)$$

where $\mathbf{E}$ is the electric field induced by the changing magnetic field and μ the permeability of free space. Since the rate of change of $\mathbf{j}$ may also be ascribed to radiation and the conduction current is carried predominantly by the more mobile electrons, the governing equation becomes

$$-\frac{e}{m} \mathbf{j} \wedge \mathbf{B} + \nu \mathbf{j} = \frac{ne^2}{m}(\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) - \frac{ne}{(n+n_i)m}\frac{\partial p}{\partial \mathbf{r}}, \qquad \ldots \quad (2.9)$$

where m is the mass of an electron and ν the mean collision frequency (Westfold 1953), which is effectively 4/3 times the electron collision frequency. Apart from the interpretation of ν, eqn. (2.9) is equivalent to the free-path eqn. (12) of Cowling (1953, p. 537). When the pressure–gradient term is negligible it is equivalent to the tensor conductivity relation most recently used by Piddington (1954 a, b, 1955)‡. In the set of equations just written down, viscosity and thermal conductivity have been neglected in accordance with a consistent scheme of approximations to Boltzmann's equation and the transport equations of kinetic

---

† The equations of this paper are expressed in terms of rationalized electromagnetic units.

‡ It seems preferable to proceed from this equation direct, rather than to introduce conductivity components which depend on quantities such as ρ and $\mathbf{B}$, which are usually variable.

theory. Their effects would need to be considered in determining the nature of the shock transition, as in Sen (1956).

Substitution from (2.2) and (2.5) into (2.4) gives

$$\tfrac{1}{2}N\frac{D \ln T}{Dt}=\frac{D \ln \rho}{Dt},$$

whence

$$\rho T^{-N/2}=\text{const.}\quad \cdots \cdots \cdots \quad (2.10)$$

in the moving gas. Again, since the gas is Lorentzian, the density is effectively given by

$$\rho=n_i m_i, \qquad \cdots \cdots \cdots \quad (2.11)$$

whence, with (2.1) and (2.5), it is not difficult to express the pressure in terms of ρ and T,

$$p=-\frac{e_i-e}{em_i}\,\rho kT, \qquad \cdots \cdots \quad (2.12)$$

and hence obtain the adiabatic equation of state

$$p\rho^{-(N+2)/N}=\text{const.}\quad \cdots \cdots \cdots \quad (2.13)$$

Before proceeding with the elimination of **E** and **j** we make a simplification of (2.9) to suit coronal conditions. First consider the two terms involving **B**. The ratio of their magnitudes is j/nev, which by (2.7) is of magnitude $B/\mu nevD$, where D is a distance within which B alters by an appreciable fraction of itself. Thus

$$\frac{j}{nev}\simeq\frac{\omega_B}{\mu\epsilon\omega_0{}^2vD}=\frac{c}{v}\frac{\omega_B}{\omega_0}\frac{c}{\omega_0 D},†$$

where $\omega_B=eB/m$ is the angular electron gyro-frequency and $\omega_0=\sqrt{(ne^2/m\epsilon)}$ is the angular electron plasma frequency, ϵ being the permittivity of free space and $c=1/\sqrt{(\mu\epsilon)}$ the velocity of electromagnetic waves in free space. Now if B is expressed in gauss (see § 5) $\omega_B=1\cdot8\times10^7B$ sec^{-1}; in the corona $\omega_0\simeq10^8$ sec^{-1} and a suitable dimension associated with large sunspots and flares is $D\simeq10^9$ cm. Thus the ratio $j/nev \simeq 10^3B/v$, which can be regarded as quite negligible in the present circumstances. It diminishes further in the chromosphere by a factor of up to 10^4. Again, the ratio of the magnitudes of the two terms on the left side is ω_B/ν. In the corona $\nu\simeq1$ sec^{-1} so that the collisional damping term, which is responsible for the ' ordinary ' electric conductivity, is clearly negligible. However, in the absence of an imposed magnetic field and below the corona it may assume importance.

Hence, in the corona, (2.9) reduces to

$$(n+n_i)e(\mathbf{E}+\mathbf{v}\wedge\mathbf{B})=\frac{\partial p}{\partial \mathbf{r}},$$

which indicates that the effective electric force on the conduction electrons is balanced by their pressure gradient. Substituting for the number

---

† We use the symbol $\simeq$ to indicate agreement to within a factor of about 10.

densities in terms of the mass density by means of (2.1) and (2.11), we get

$$-\frac{e_i - e}{m_i}\,\rho(\mathbf{E} + \mathbf{v} \wedge \mathbf{B}) = \frac{\partial p}{\partial \mathbf{r}} \qquad \ldots \ldots \quad (2.14)$$

and, using this result to eliminate $\mathbf{E}$ from (2.6),

$$-\frac{\partial}{\partial \mathbf{r}} \wedge (\mathbf{v} \wedge \mathbf{B}) - \frac{m_i}{e_i - e}\,\frac{\partial}{\partial \mathbf{r}} \wedge \left(\frac{1}{\rho}\,\frac{\partial p}{\partial \mathbf{r}}\right) + \frac{\partial \mathbf{B}}{\partial t} = 0.$$

But

$$\frac{\partial}{\partial \mathbf{r}} \wedge \left(\frac{1}{\rho}\,\frac{\partial p}{\partial \mathbf{r}}\right) = \frac{\partial 1/\rho}{\partial \mathbf{r}} \wedge \frac{\partial p}{\partial \mathbf{r}}$$

$$= 0$$

since by (2.13) $\partial p/\partial \mathbf{r}$ and $\partial \rho/\partial \mathbf{r}$ are parallel. Hence the rate of change of magnetic induction is simply given by

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{\partial}{\partial \mathbf{r}} \wedge (\mathbf{v} \wedge \mathbf{B}), \qquad \ldots \ldots \quad (2.15)$$

which indicates that the magnetic field is effectively 'frozen' into the medium, as if the conductivity were infinite. Where the collisional term in (2.9) is not negligible, the dissipative term†

$$\frac{\nu m m_i}{\mu e e_i}\,\frac{\partial}{\partial \mathbf{r}} \wedge \left(\frac{1}{\rho}\,\frac{\partial}{\partial \mathbf{r}} \wedge \mathbf{B}\right)$$

should be added to the right side of (2.15). An alternative form of (2.15) is

$$\frac{D\mathbf{B}}{Dt} + \mathbf{B}\,\frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{v} = \mathbf{B} \cdot \frac{\partial}{\partial \mathbf{r}}\,\mathbf{v}. \qquad \ldots \ldots \quad (2.16)$$

Again, to eliminate $\mathbf{j}$ from (2.3), we have, by (2.7), the magnetic force density

$$\mathbf{j} \wedge \mathbf{B} = \frac{1}{\mu}\left(\frac{\partial}{\partial \mathbf{r}} \wedge \mathbf{B}\right) \wedge \mathbf{B}$$

$$= \frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{S},$$

where

$$\mathbf{S} = \frac{1}{\mu}\,(\mathbf{B}\mathbf{B} - \tfrac{1}{2}B^2\mathbf{U}) \qquad \ldots \ldots \quad (2.17)$$

is the magnetic stress tensor and $\mathbf{U}$ the unit tensor. Thus

$$\rho\,\frac{D\mathbf{v}}{Dt} = \frac{\partial}{\partial \mathbf{r}} \cdot (\mathbf{S} - p\mathbf{U}), \qquad \ldots \ldots \quad (2.18)$$

whence the motion may be regarded as due jointly to a magnetic stress and a hydrostatic pressure system.

---

† It follows from (2.10) and the result (cf. Westfold 1953) $\nu \propto \rho T^{-3/2}$ that ν is constant for the adiabatic changes considered here.

The quantities ρ, **v**, p and **B** are now determined by the set of first-order eqns. (2.2), (2.13), (2.16), (2.17) and (2.18) ; then **j**, **E** and T will be given by (2.7), (2.14) and (2.12).

The details of the shock transition depend on the conservation of the momentum density $\rho \mathbf{v}$ and the total energy density $\frac{1}{2}\rho v^2 + I$, where

$$I = U + B^2/2\mu \qquad \ldots \quad \ldots \quad (2.19)$$

may be called the *intrinsic energy density* of the gas ; it consists of the internal (molecular) energy density

$$U = \tfrac{1}{2}(n + n_i)NkT \qquad \ldots \quad \ldots \quad (2.20)$$

and the magnetic energy density $B^2/2\mu$. The equations governing these quantities are readily obtained from (2.2), (2.18), (2.4) and (2.16). We give only the equation for the total energy density

$$\frac{D}{Dt}(\tfrac{1}{2}\rho v^2 + I) + (\tfrac{1}{2}\rho v^2 + I)\frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{v} = \frac{\partial}{\partial \mathbf{r}} \cdot \{(\boldsymbol{S} - p\boldsymbol{U}) \cdot \mathbf{v}\}, \qquad (2.21)$$

which states that the rate of increase of the total energy within a moving element of the gas is equal to the rate of working of the total stress system on the boundaries of the element.

It should be noted that the formulae (2.5) and (2.20) directly relate U and p,

$$U = \tfrac{1}{2}Np. \qquad \ldots \quad \ldots \quad \ldots \quad (2.22)$$

§ 3. Small Disturbances in Uniform Steady Rectilinear Flow

At the present stage it is sufficient to investigate the simplest flows that will yield results relevant to the physical situation under discussion. The remainder of this paper will therefore be concerned with rectilinear flow in the direction of the coordinate axis Ox, in which all physical quantities depend only on x and t. We denote longitudinal and transverse components by the subscripts L and T. Then the eqn. (2.8) gives

$$\frac{\partial E_{\mathrm{L}}}{\partial x} = \frac{\partial B_{\mathrm{L}}}{\partial x} = 0,$$

so that the longitudinal component of (2.16) gives

$$\frac{\partial B_{\mathrm{L}}}{\partial t} = 0,$$

whence we conclude that

$$B_{\mathrm{L}} = \mathrm{const.}$$

Again, the transverse component of (2.18) with (2.17) gives

$$B_{\mathrm{L}}\frac{\partial B_{\mathrm{T}}}{\partial x} = 0.$$

Of the alternatives presented by this result, the case where $\partial B_{\mathrm{T}}/\partial x = 0$ is that of a motion independent of the magnetic field, which might be associated with a Type II burst ; the corresponding hydrodynamic

shock conditions are well known. Accordingly, we take **B** as perpendicular to Ox so that the eqns. (2.2), (2.16), (2.17) and (2.18) reduce to

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho v) = 0, \qquad \cdots \quad (3.1)$$

$$\frac{\partial B}{\partial t} + \frac{\partial}{\partial x}(Bv) = 0, \qquad \cdots \quad (3.2)$$

$$\rho\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial x}\right) = -\frac{\partial}{\partial x}\left(p + \frac{B^2}{2\mu}\right), \qquad \cdots \quad (3.3)$$

which, with (2.13), determine ρ, v, p and B. It follows, too, from (2.7) that **j** is also transverse to the direction of motion and perpendicular to **B**. It will be observed that only the isotropic part of the magnetic stress tensor contributes to (3.3), augmenting the hydrostatic pressure p with the magnetic pressure $B^2/2\mu$. The similarity between (3.1) and (3.2) is due to the 'freezing' into the gas of the field lines.

For the case of steady flow (3.1) and (3.2) provide continuity integrals

$$\rho v = \text{const.}, \qquad \cdots \quad (3.4)$$

$$Bv = \text{const.}, \qquad \cdots \quad (3.5)$$

with which (3.3) gives the momentum integral

$$\rho v^2 + p + \frac{B^2}{2\mu} = \text{const.} \qquad \cdots \quad (3.6)$$

In the same way, (2.21) with (2.17) and (3.4) yields the Bernoulli integral

$$\tfrac{1}{2}v^2 + \left\{\left(U + \frac{B^2}{2\mu}\right) + \left(p + \frac{B^2}{2\mu}\right)\right\}\bigg/\rho = \text{const.} \qquad \cdots \quad (3.7)$$

The second term on the left side represents the specific enthalpy of the gas in the magnetic field.

In order to consider the propagation of small disturbances in uniform steady flow we replace the variables ρ, v, p and B in (3.1)–(3.3) by $\rho_0 + \rho$, etc., where the subscript 0 denotes constant steady-state values, and retain only first-order terms in ρ, etc. We then seek solutions which are functions of a new variable

$$\xi = x - Ct \qquad \cdots \quad (3.8)$$

where the propagation velocity C is to be determined. The equations in terms of ξ are

$$(C - v_0)\frac{d\rho}{d\xi} = \rho_0\frac{dv}{d\xi}, \qquad \cdots \quad (3.9)$$

$$(C - v_0)\frac{dB}{d\xi} = B_0\frac{dv}{d\xi}, \qquad \cdots \quad (3.10)$$

$$(C - v_0)\frac{dv}{d\xi} = \frac{a_0^2}{\rho_0}\frac{d\rho}{d\xi} + \frac{b_0^2}{B_0}\frac{dB}{d\xi}, \qquad \cdots \quad (3.11)$$

where

$$a_0^2 = \left(\frac{dp}{d\rho}\right)_0, \qquad b_0^2 = \frac{B_0^2}{\mu\rho_0}. \qquad \cdots \quad (3.12)$$

The quantity a_0 is the usual sound velocity in the absence of a magnetic field and b_0 is Alfvén's magnetohydrodynamic velocity in an incompressible gas. By (2.13)

$$a_0{}^2 = \frac{N+2}{N} \frac{p_0}{\rho_0} . \qquad \cdots \cdots \quad (3.13)$$

Elimination of the derivatives from (3.9)–(3.11) now gives the equation

$$(C-v_0)^2 = a_0{}^2 + b_0{}^2, \qquad \cdots \cdots \quad (3.14)$$

which determines C. Thus, any small disturbance will in general be propagated in both directions parallel to Ox with the speed c_0, given by (1.2), relative to the stream velocity v_0. This is the formula first obtained by de Hoffmann and Teller (1950), who also point out that these waves are distinct from Alfvén waves which are transverse and travel in directions parallel to $\mathbf{B}_0$. In § 4 we shall see that c_0 provides a first approximation to the velocity of a shock front at the head of a corpuscular stream in a magnetic field in the solar corona.

§ 4. The Shock Velocity and Jump Conditions

The velocity of a shock front and the jump conditions across it can be derived by considering the conservation of mass, of the magnetic field lines and of the momentum and energy of a column of moving gas containing a discontinuity, as in Courant and Friedrichs (1948, § 54). The results are the same as those obtained by Lüst (1953) for the case where the flow relative to a coordinate system moving with the shock front is steady. For the present case of perpendicular longitudinal shock, where the front moves with the velocity V and the relative velocities on both sides are

$$V_1 = v_1 - V, \qquad V_0 = v_0 - V, \qquad \cdots \quad (4.1)$$

the eqns. (3.1)–(3.3) and (2.21) yield the results

$$\rho_1 V_1 = \rho_0 V_0 = M, \qquad \cdots \cdots \quad (4.2)$$

$$B_1 V_1 = B_0 V_0 = h, \qquad \cdots \cdots \quad (4.3)$$

$$\rho_1 V_1{}^2 + P_1 = \rho_0 V_0{}^2 + P_0 = \Pi, \qquad \cdots \cdots \quad (4.4)$$

where

$$P_1 = p_1 + B_1{}^2/2\mu, \qquad P_0 = p_0 + B_0{}^2/2\mu, \qquad \cdots \quad (4.5)$$

and the Rankine–Hugoniot condition

$$\tfrac{1}{2} V_1{}^2 + (I_1 + P_1)/\rho_1 = \tfrac{1}{2} V_0{}^2 + (I_0 + P_0)/\rho_0 = \tfrac{1}{2} q^2. \qquad \cdots \quad (4.6)$$

In these, M, h and Π are the flux densities through the front of the mass, the field-line intensity and the total momentum, and q is the relative limit speed. The relations (4.2)–(4.6) have the same form as the steady-state integrals (3.4)–(3.7). The specifically magnetic contributions to the magnetohydrodynamic shock conditions are embodied in (4.3) and the extended definitions (2.19) and (4.5) of energy density and pressure.

For the present purpose we envisage a stream of ionized gas moving with a velocity v_1 behind a shock front advancing into a similar medium of given density ρ_0 at rest at a given pressure p_0 in a given field of magnetic induction B_0. From the four jump conditions just obtained and the relation (2.22) we seek the velocity V of the front and the density ρ_1, pressure p_1 and induction B_1 behind the front. Although the algebra would be simplified by taking $N=3$ at the outset, we retain the general value in order to facilitate comparison with the standard hydrodynamic results for polytropic gases, and to exhibit the nature of the dependence of the results on the value adopted.

We proceed by eliminating ρ_1 and B_1 from (4.4) and (4.6) by means of (4.2) and (4.13); then, elimination of p_1 gives a relation between V_1 and V_0 which, by (4.1) determines V. Hence V_1 and V_0 are known, so that (4.2) and (4.3) determine ρ_1 and B_1; finally p_1 is determined by (4.4).

The first step gives

$$M V_1 + p_1 + h^2/2\mu V_1{}^2 = \Pi,$$
$$M V_1 + (N+2)p_1 + 2h^2/\mu V_1{}^2 = Mq^2/V_1,$$

whence

$$(N+1)M V_1 + (N-2)h^2/2\mu V_1{}^2 = (N+2)\Pi - Mq^2/V_1.$$

Substitution of the values of M, Π and q^2 then gives

$$[\rho_0 V_1 V_0\{(N+1)V_1 - V_0\} - (B_0{}^2/2\mu)\{(N+2)V_1 + (N-2)V_0\}$$
$$-(N+2)p_0 V_1](V_1 - V_0) = 0,$$

or, since $V_1 \neq V_0$, on substitution from (3.12) and (3.13),

$$V_1 V_0\{(N+1)V_1 - V_0\} - \tfrac{1}{2}b_0{}^2\{(N+2)V_1 + (N-2)V_0\} - Na_0{}^2 V_1 = 0. \qquad (4.7)$$

In the absence of a magnetic field, $b_0 = 0$ and (4.7) agrees with the quadratic relation (17.08) of Courant and Friedrichs (1948) for a polytropic gas, and yields the result (1.1) when we take $N=3$. In the present case, however, substitution from (4.1) gives a cubic equation which determines the shock speed V.

The solution of the cubic is facilitated by taking $v_0 = 0$, as in the coronal situation that we envisage. This procedure will involve no loss of generality as the results will be extended to the case of a gas in motion on both sides of the front, by means of a Galilean transformation. Then (4.7) gives the cubic

$$V(V-v_1)\{NV - (N+1)v_1\} - \tfrac{1}{2}b_0{}^2\{2NV - (N+2)v_1\} - Na_0{}^2(V-v_1) = 0.$$
$$\qquad \cdot \quad \cdot \quad \cdot \quad (4.8)$$

Since under coronal conditions we may expect $v_1 \ll V$, the corresponding first approximation to V is easily found to be the generalized sound velocity c_0. The next approximation

$$V = c_0 + \frac{2(N+1)a_0{}^2 + 3Nb_0{}^2}{4Nc_0{}^2} v_1 \qquad \cdot \quad \cdot \quad \cdot \quad (4.9)$$

is sufficient for our present purpose. The fact that the shock velocity is little more than the sound velocity in this case indicates that streams of 'auroral' particles in the corona give rise to only weak magnetohydrodynamic shocks.

The corresponding velocities relative to the front are now given by (4.1), and are valid for any Galilean frame of reference if v_1 is replaced by $v_1 - v_0$. Thus, if (4.9) is written in the abbreviated form

$$V = c_0 + \alpha v_1,$$

relative to any Galilean frame we should have

$$-V_0 = c_0 + \alpha(v_1 - v_0),$$
$$-V_1 = c_0 - (1 - \alpha)(v_1 - v_0),$$

values which are respectively supersonic and subsonic. Then, by (4.2) and (4.3),

$$\frac{\rho_1}{\rho_0} = \frac{B_1}{B_0} = \frac{V_0}{V_1} = 1 + \frac{v_1 - v_0}{c_0}$$

approximately, whence the jumps are given by

$$\frac{\rho_1 - \rho_0}{\rho_0} = \frac{B_1 - B_0}{B_0} = \frac{v_1 - v_0}{c_0}, \qquad \cdots \cdots \text{ (4.10)}$$

which is independent of N. It follows from (2.7) that this discontinuity in B will give rise to a surface current in the shock front of strength $\mathbf{n} \wedge (\mathbf{B}_1 - \mathbf{B}_0)/\mu$ per unit area, where $\mathbf{n}$ is the unit vector in the direction of the flow. Again, the pressure jump is determined by (4.4). which, with (4.5) and (4.10) gives

$$p_1 - p_0 = -\left(\rho_0 V_0 c_0 + \frac{B_0^2}{\mu}\right)\frac{v_1 - v_0}{c_0} = 0,$$

whence, on replacing V_0 by $-c_0$ and substituting from (3.12) and (3.13). we get for the excess pressure ratio

$$\frac{p_1 - p_0}{p_0} = \frac{N+2}{N}\frac{v_1 - v_0}{c_0}, \qquad \cdots \cdots \text{ (4.11)}$$

to the same order of approximation. Again, the magnitude of these jumps is characteristic of a weak shock. Thus, at the shock front, both $\mathbf{j}$ and $\partial p/\partial \mathbf{r}$ are represented by delta functions. There, the first term on the left side of (2.9) is not negligible. It is then not difficult to show that there must be a surface distribution of electric dipole moment of strength

$$\epsilon m_i \left(\frac{a_0^2}{e_i - e} + \frac{b_0^2}{e_i}\right)\frac{v_1 - v_0}{c_0}\,\mathbf{n},$$

per unit area.

There remains the determination of the jumps in the quantities T, j and E, which were previously eliminated. The first is obtained by substitution from (4.10) and (4.11) into (2.12), whence

$$\frac{T_1 - T_0}{T_0} = \frac{2}{N}\frac{v_1 - v_0}{c_0}. \qquad \cdots \cdots \text{ (4.12)}$$

Since, by (2.7) and (2.14), the jumps in j and E involve the jumps in the derivatives of B and p, respectively, these cannot be determined without investigation of a specific flow problem over a range of x on both sides of the shock front. However, if, as may be expected in the present case, the relative jump in $\partial p/\partial r$ is no more than $O((v_1-v_0)/c_0)$, the principal part of the jump in $\mathbf{E}$ will be due to the jump in the magnetic force acting on the medium, so that

$$\mathbf{E}_1 - \mathbf{E}_0 \doteqdot -(\mathbf{v}_1-\mathbf{v}_0) \wedge \mathbf{B}_0, \qquad \dots \quad (4.13)$$

which is more substantial in magnitude than those above.

It is of interest to give the corresponding jump relations for a strong hydrodynamic shock such as we have associated with a Type II burst. Proceeding from (4.8) with $b_0=0$, we get

$$\left.\begin{aligned} -V_0 &= \frac{N+1}{2N}(v_1-v_0) + \sqrt{\left\{\left(\frac{N+1}{2N}\right)^2 (v_1-v_0)^2 + a_0^2\right\}}, \\ -V_1 &= -\frac{N-1}{2N}(v_1-v_0) + \sqrt{\left\{\left(\frac{N+1}{2N}\right)^2 (v_1-v_0)^2 + a_0^2\right\}}. \end{aligned}\right\} \quad (4.14)$$

Then, by (4.2), the condensation

$$\frac{\rho_1-\rho_0}{\rho_0} = -\frac{v_1-v_0}{V_1}. \qquad \dots \quad (4.15)$$

Again, by (4.4) and (3.13), the excess pressure ratio

$$\frac{p_1-p_0}{p_0} = -\frac{N+2}{N}\frac{V_0(v_1-v_0)}{a_0^2}, \qquad \dots \quad (4.16)$$

and finally, by (2.12) with (3.13),

$$\frac{T_1-T_0}{T_0} = \frac{a_1^2-a_0^2}{a_0^2}. \qquad \dots \quad (4.17)$$

The jump in E is again indeterminate with the present data.

§ 5. Numerical Discussion

For our present purpose, the important quantities to estimate are the velocity V of the shock front and the magnitudes of the jumps which, in the magnetohydrodynamic case, are all, save for the electric field E, of relative magnitude $O(v_1/c_0)$.

By (3.13) and (2.12),

$$a_0^2 = -\frac{N+2}{N}\frac{e_i-e}{em_i}kT_0. \qquad \dots \quad (5.1)$$

Since the corona consists predominantly of fully ionized hydrogen, we may take $e_i=-e$, which, with $N=3$, gives

$$a_0^2 = 2 \cdot 76 \times 10^8 \, T_0 (\text{cm/sec})^2.$$

Now all physical estimates agree that T_0 is about 10^6 °K, whence we find that

$$a_0 = 170 \text{ km/sec},$$

In solar physics, magnetic induction fields are measured in gauss, which are c.g.s. electromagnetic units ; in the present rationalized system $\mu=4\pi$, so that the second formula (3.12) gives

$$b_0{}^2 = \frac{B_0{}^2}{4\pi\rho_0} . \qquad\qquad \dots \dots \dots \quad (5.2)$$

For $e_i=-e$ (2.11) and (2.1) give $\rho=nm_i$, whence

$$b_0{}^2 = 4 \cdot 8 \times 10^{22} B_0{}^2/n_0 \ (\text{cm/sec})^2.$$

In the lower corona n_0 is about 10^8 cm^{-3}, whence

$$b_0 = 2 \cdot 2 \times 10^2 B_0 \ \text{km/sec},$$

for B_0 given in gauss.

If now we identify the velocity of a Type III burst with the velocity of a magnetohydrodynamic shock wave ahead of a corpuscular stream moving in a magnetic field, we can obtain an estimate of the magnitude if the field required. As a good approximation we may neglect both a_0 and v_1 as compared to b_0, so that (4.9) and (1.2) give

$$V \doteqdot c_0 \doteqdot b_0.$$

Thence we find that shock velocities between 3×10^4 and 2×10^5 km/sec require coronal magnetic fields of between about 150 and 1000 gauss. If (4.13) is valid, the jump in electric field across the shock front ahead of a stream moving at 1000 km/sec would, for the above values of the magnetic field, lie between about 150 and 1000 volt/cm.

On the other hand, the magnetohydrodynamic shock velocity corresponding to a stream velocity of 1000 km/sec is 1400 km/sec, as was shown in § 1. In this case (4.15) and (4.16) give for the condensation and excess pressure ratio

$$(\rho_1-\rho_0)/\rho_0 = 2 \cdot 8 \quad \text{and} \quad (p_1-p_0)/p_0 = 78.$$

Thus $a_1{}^2/a_0{}^2 = 21$ so that (4.17) gives

$$(T_1-T_0)/T_0 = 20.$$

It should be pointed out that these very large values are considerably reduced if the shock velocity is taken as 500 km/sec, the estimated velocity of a typical Type II burst. Then the associated stream velocity is 330 km/sec, whence

$$(\rho_1-\rho_0)/\rho_0 = 1 \cdot 9, \quad (p_1-p_0)/p_0 = 9 \cdot 5$$

and

$$(T_1-T_0)/T_0 = 2 \cdot 6.$$

An essential requirement of the present interpretation of Type III burst velocities as those of magnetohydrodynamic shock waves is that large magnetic fields, of some hundreds of gauss, are at the same time present in the corona. Measurements at the level of the photosphere do give such values, but, on the basis of the currently proposed static models of sunspot fields, it is difficult to see how such fields could be maintained in the corona. While there are no observational techniques at present capable of magnetic-field measurements that could determine the matter, one can only point

out that the phenomena under discussion are essentially *dynamic*, and indicate the kind of dynamic situation which might produce the required conditions. If we accept Allen's (1944) conclusion that coronal streams and *M* regions are probably identical, and, with Kiepenheuer (1953, § 5.4), suppose that the streams originate in explosions connected with the development of filaments in the chromosphere below, we may expect magnetic field lines to be convected up from the photosphere with any moving material within the filament, and to maintain the same intensity so long as the material density remains the same.

Such motions are observed in prominences in the neighbourhood of sunspots. Continuity considerations and the magnetic jump condition indicate that there must be a comparable field in the corona at the time of an explosion in such filamentary material, which, as new values of velocity, density and magnetic field are impressed on the medium behind the advancing shock front, will effectively blow out a segment of the corona. The details must await the investigation of a model situation, but it appears that, because of its non-uniformity, the medium ahead of the front will not remain at rest after the explosion, nor will the density and magnetic-field distributions retain their initial values. It may be expected that these values will be augmented, to a negligible extent at remote distances, and more considerably in the neighbourhood of the front.

As indicated at the beginning of § 2, it is suggested that radiation will result from perturbation of the values of the quantities specifying the state of the gas as hitherto considered. If perturbed values are specified at the shock front, the resulting transient field will be determined. The situation is of a similar character to that considered by Jaeger and Westfold (1949), with the additional complications occasioned by the moving boundary.

References

Allen, C. W., 1944, *Mon. Not. R. astr. Soc.*, **104**, 13.
Chapman, S., and Bartels, G., 1940, *Geomagnetism* (Oxford : Clarendon Press).
Courant, R., and Friedrichs, K. O., 1948, *Supersonic Flow and Shock Waves* (New York : Interscience Publishers).
Cowling, T. G., 1953, *The Solar System*, Vol. 1, The Sun, edited by G. P. Kuiper (Chicago : University Press).
de Hoffmann, F., and Teller, E., 1950, *Phys. Rev.*, **80**, 692.
Jaeger, J. C., and Westfold, K. C., 1949, *Aust. J. sci. Res.* A, **2**, 322.
Kiepenheuer, K. O., 1953, *The Solar System*, Vol. 1, The Sun, edited by G. P. Kuiper (Chicago : University Press).
Lüst, R., 1953, *Z. Naturf.*, **8 a**, 277.

PAWSEY, J. L., and BRACEWELL, R. N., 1955, *Radio Astronomy* (Oxford : Clarendon Press).
PAYNE-SCOTT, R., and LITTLE, A. G., 1952, *Aust. J. sci. Res.* A, **5,** 32.
PIDDINGTON, J. H., 1954 a, *Mon. Not. R. astr. Soc.*, **114,** 638 ; 1954 b, *Ibid.*, **114,** 651 ; 1955, *Ibid.*, **115,** 671.
SEN, H. K., 1956, *Phys. Rev.*, **102, 5**.
WESTFOLD, K. C., 1953, *Phil. Mag.*, **44, 711**.
WILD, J. P., 1950 a, *Aust. J. sci. Res.* A, **3, 399** ; 1950 b, *Ibid.*, **3,** 541.
WILD, J. P., and MCCREADY, L. L., 1950, *Aust. J. sci. Res.* A, **3,** 387.
WILD, J. P., MURRAY, J. D., and ROWE, W. C., 1953, *Nature, Lond.*, **172,** 533.
WILD, J. P., ROBERTS, J. A., and MURRAY, J. D., 1954, *Nature, Lond.*, **173,** 532.

Electron Diffraction from Crystals Containing Stacking Faults: II†

By M. J. Whelan and P. B. Hirsch
Crystallographic Laboratory, Cavendish Laboratory,
Cambridge, England

Abstract

The theory of the previous paper is applied to the case of stacking faults in close packed lattices. Electron optical experiments on thin foils of stainless steel are described, the results of which are in good agreement with theory.

§ 1. Introduction

In the course of transmission microscopy experiments on dislocations in austenitic (face-centred cubic) stainless steel foils (Bollmann 1957, Whelan *et al.* 1957) certain interference fringes were observed which were attributed to stacking faults produced on (111) slip planes. In the present paper these interference effects are discussed in detail and interpreted in terms of the theory developed in the previous paper (Whelan and Hirsch 1957, to be referred to as I). Both the nature of the fringes on micrographs and that of the diffraction patterns are considered. The general agreement obtained between theory and experiment confirms that the fringes are due to stacking faults and that the theory developed in I is basically correct.

In I the dynamical theory was developed for a plate-like crystal containing a stacking fault characterized by an arbitrary shear **R**, which gives rise to a phase difference $\alpha = 2\pi\mathbf{g}\,.\,\mathbf{R}$ in the electron waves diffracted from either side of the fault. Here the theory is applied to the particular case of a stacking fault in a close packed lattice. For convenience the predictions of the theory are discussed first, and are followed by a detailed interpretation of the experimental observations on stainless steel.

§ 2. Stacking Faults in Close Packed Lattices

It is well known that the stacking of close packed planes of atoms (fig. 1) can be achieved in a large number of ways without violating nearest neighbour interactions owing to the fact that each successive layer of atoms can be added in two ways. In fig. 1 each atom A is surrounded by six hollows, only three of which are occupied by atoms

---

† Communicated by the Authors.

of the layer above, which may be placed in either the B or C positions. Of the possible arrangements only two, namely the face-centred cubic and hexagonal close packed structures, lead to crystallographically repeating structures. In the face-centred cubic structure the stacking sequence follows the pattern ABCABCABC, while in the hexagonal structure the sequence is ABABABAB. All other arrangements may be described in terms of these ideal structures together with suitable distributions of stacking faults. A stacking fault occurs when there is a

Fig. 1
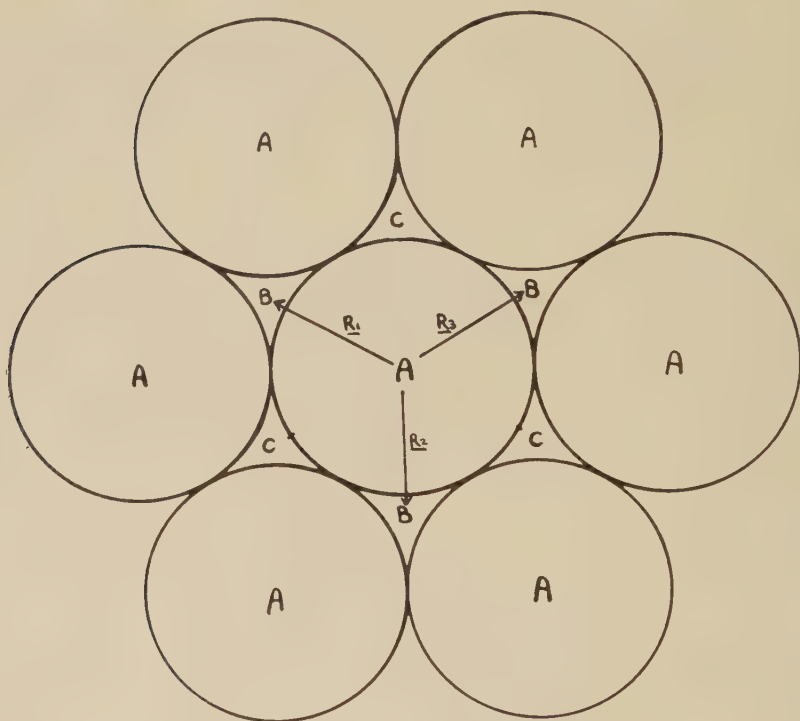
Stacking sequence in the face-centred cubic lattice. Atoms A in the plane of the paper ; the next layer has atoms in the B positions ; the third layer has atoms in the C positions. The shear vectors of a stacking fault are indicated on the figure.

mistake in the stacking sequence such as ABCABC BCABC. The types of stacking faults in close packed lattices have been discussed by Frank (1951), who classifies them into intrinsic and extrinsic faults. The above is an example of an intrinsic fault ; it may be produced by shearing the top half of the crystal in fig. 1 relative to the bottom half by any of the vectors R_1, R_2, R_3, or by removal of a close packed layer with the subsequent closing of the gap. This type of fault is also

produced when partial dislocations separate. For an extrinsic fault the stacking sequence is ABCABC B ABCABC ; it may be produced by the insertion of an extra layer of atoms, or by two intrinsic faults on adjacent (111) planes. In this paper interest is centred mainly on intrinsic faults, since this type of fault is common to both cubic and hexagonal close packed lattices, and as far as this theory is concerned, all other types of faults may be considered as a suitable distribution of intrinsic faults.

Fig. 2
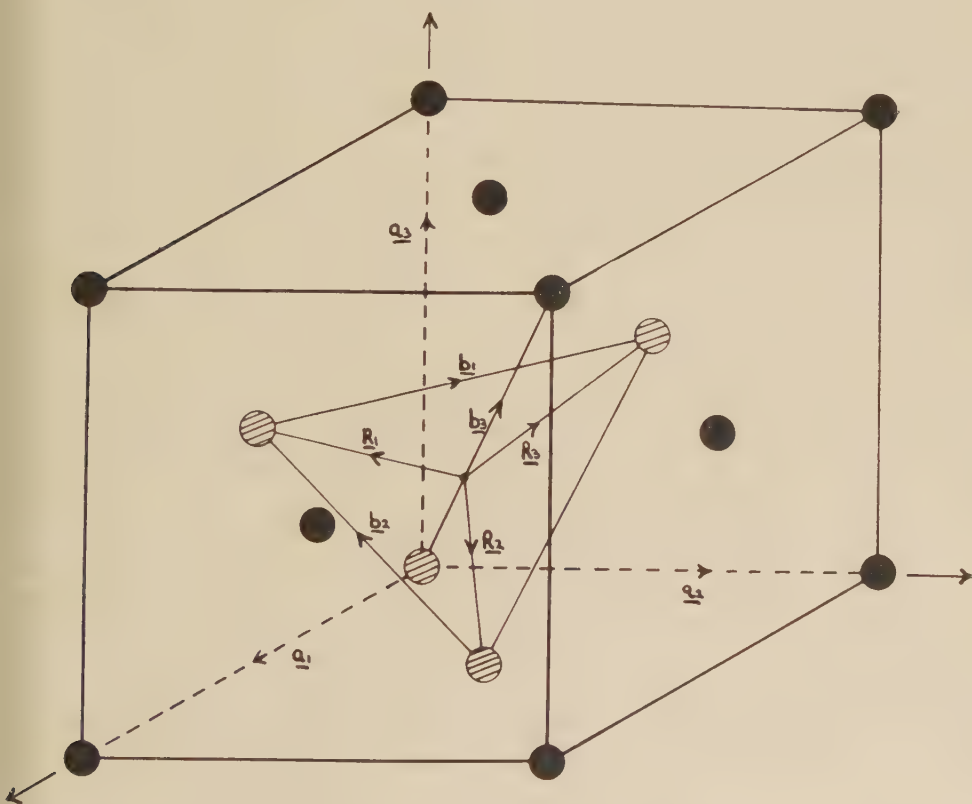
Face-centred cubic structure showing shear vectors $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$ for an intrinsic stacking fault on (111). The choice of cubic, $\mathbf{a}_i$ and hexagonal, $\mathbf{b}_i$, axes is indicated.

A stacking fault is therefore produced by any one of the vector displacements $\mathbf{R}_1$, $\mathbf{R}_2$, $\mathbf{R}_3$ (figs. 1 and 2), and except for the effect on the external shape of the crystal all three translations are equivalent, so that $\mathbf{R}_1$ only need be considered. Using cubic axes $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ as in fig. 2, $\mathbf{R}_1 = (a/6)[1\bar{2}1]$ for a stacking fault on (111). Putting $\mathbf{g} = (h, k, l)$, it is then seen that $\alpha = \pi/3\{h - 2k + l\}$. It is useful at this point to transform

to the hexagonal axes $\mathbf{b}_i$ shown in fig. 2. The relation between the hexagonal indices (HKL) and the cubic indices (hkl) is as follows
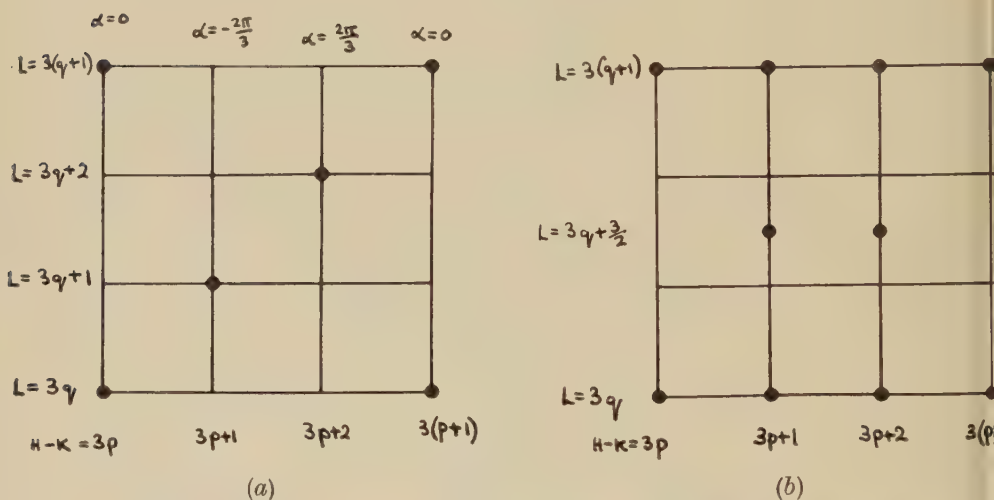
$$H=-\tfrac{1}{2}h+\tfrac{1}{2}k$$
$$K=-\tfrac{1}{2}k+\tfrac{1}{2}l$$
$$L=(h+k+l).$$

It follows that $(h-2k+l)=-2(H-K)$ and that $\alpha=-2\pi(H-K)/3$. Remembering that the face-centred cubic reflections are given by hkl

Fig. 3

(a) Reflections from a face-centred cubic lattice indexed on the hexagonal system.
(b) Reflections from an hexagonal lattice corresponding to fig. 3(a).

either all odd or all even, examination of the above expressions for H, K and L show that the reflections belong to the following groups

$$L=3q, \ (H-K)=3p, \qquad\qquad \alpha=0$$
$$L=3q+1, \ (H-K)=3p+1, \qquad \alpha=-2\pi/3$$
$$L=3q+2, \ (H-K)=3p+2, \qquad \alpha=+2\pi/3$$

where p and q are integers. In these expressions the principal value of α lying between $-\pi$ and $+\pi$ has been chosen. The body-centred reciprocal lattice for a face-centred cubic structure can therefore be regarded as a stacking of hexagonal layers of points parallel to the faulting plane repeating every third layer as in fig. 3 (a). The phase angle α is then either 0, $+2\pi/3$, or $-2\pi/3$ depending on the layer in which the reciprocal lattice point (hkl) is situated. A similar interpretation was used by Paterson (1952) in a kinematical theory of x-ray diffraction from a crystal containing a random distribution of stacking faults. It can be seen that the reflections with zero phase factor are

simply those which are not displaced or broadened in his theory.   Further-more, since α depends only on $(H-K)$ and not on L, the reflections of a hexagonal close packed lattice (fig. 3 (b)) will have similar phase factors determined by $(H-K)$ only.

§ 3. GENERAL FRINGE PROFILE

In I expressions were derived for electron wave functions at the lower surface of the crystal, and formulae were obtained predicting interference fringes running parallel to the line of intersection of the stacking fault

Fig. 4
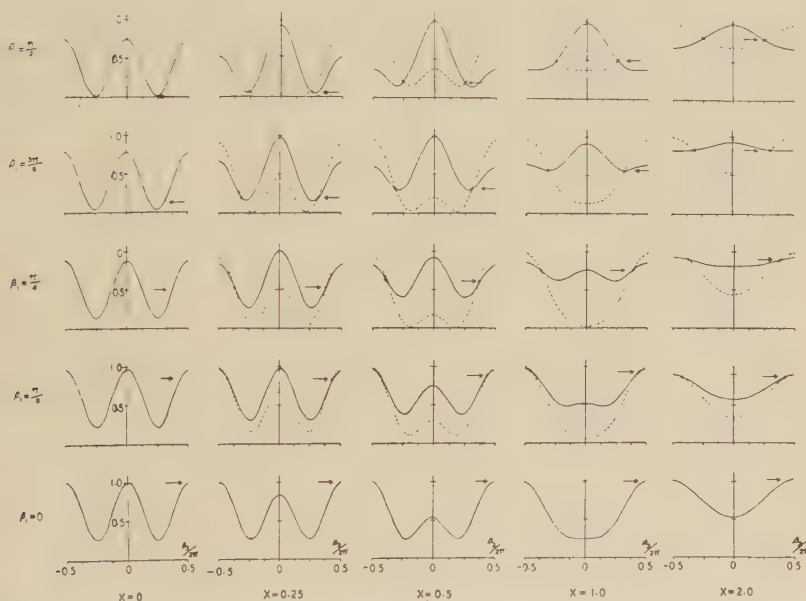
The theoretical intensity profile ζ of the fringes at a stacking fault as a function of $\beta_2 = 2\pi z \sqrt{(1+x^2)}, t_0$. The different curves correspond to various values of x and $\beta_1 = \pi t \sqrt{(1+x^2/t_0)}$. The appearance of subsidiary maxima for $|x| < 1$ is evident. For $|x| \gg 1$ the profile approximates for most values of β_1 to a simple cosine function. The arrows indicate the value of the intensity in the unfaulted regions.

with the crystal surface.   Equation (24) of I may now be applied to the case $\alpha = 2\pi/3$. Putting $|\xi_0|^2 = \zeta$ the expression for the intensity profile of the fringes becomes

$$\zeta = [\cos\beta_1 + 3(\cos\beta_2 - \cos\beta_1)/4(1+x^2)]^2$$
$$+ [\sqrt{3}(\cos\beta_2 - \cos\beta_1)/4(1+x^2) - x\sin\beta_1/\sqrt{(1+x^2)}]^2 \quad (1)$$

where

$$\beta_1 = \pi t \sqrt{(1+x^2)}, t_0, \quad \beta_2 = 2\pi z \sqrt{(1+x^2)}, t_0$$
(see notation of I).

The intensity profile may also be expressed in the form

$$\zeta = A_1 + A_2 \cos \beta_2 + A_3 \cos^2 \beta_2$$

where A_1, A_2, A_3 are functions of the parameters x and β_1 both of which may vary along the length of the fault.

Figure 4 shows how this function varies with x and β_1 in the range $0 \leqslant \beta_1 \leqslant \frac{1}{2}\pi$, $-\pi \leqslant \beta_2 \leqslant \pi$. The function is periodic in β_1, β_2 with period 2π, and may be obtained in the region $\frac{1}{2}\pi \leqslant \beta_1 \leqslant 2\pi$ by noting that

$$\zeta((\pi - \beta_1), x, \beta_2) = \zeta(\beta_1, (-x), (\pi + \beta_2))$$
$$\zeta((\pi + \beta_1), x, \beta_2) = \zeta(\beta_1, x, (\pi + \beta_2)).$$

The full and dotted curves refer to x positive and negative respectively. In general the function consists of alternate strong and weak maxima. The latter give rise to the doublet appearance of the minima characteristic of the dynamical region, shown schematically in fig. 5 (a). In the kinematical region A_3 is small compared with A_2 and as a result the subsidiary maxima are absent. The minima therefore appear essentially as singlets (fig. 5 (b)).

In discussing the experimental observations it will be necessary to consider the change in the fringe system in regions of foil where the thickness or orientation varies gradually. In the former case β_1 varies but x is constant. When $\tan \beta_1 = \sqrt{3}x / \sqrt{(1 + x^2)}$, $A_2 = 0$ and the maxima are of equal intensity. The dark and light fringes are then spaced equally at distances $\frac{1}{2} t_0' \cot \phi$ (see notation I). In general this occurs at $\beta_1 = n\pi + \epsilon$; when x is small $\epsilon \simeq 0$, and this occurs near a maximum of transmitted intensity in the unfaulted crystal. Further inspection of fig. 4 shows that the fringes are reversed on either side of such values of β_1, in the sense that weak maxima become strong maxima and *vice versa*. The way the fringes reverse as β_1 is varied is shown schematically in fig. 5 (a).

For large x, $\epsilon = \pi/3$ so that the fringe reversal takes place at a value of β_1 displaced from that corresponding to maximum intensity in the unfaulted crystal. Further, since A_3 is negligible and $A_2 = 0$, the fringes are very weak in this case, and the fringe reversal will appear as shown schematically in fig. 5 (b).

With regard to gradual changes of orientation of the foil, as x is decreased the fringe separation $t_0' = t_0 / \sqrt{(1 + x^2)}$ increases, particularly for $|x| > 1$, on going from the kinematical to the dynamical regions. Further, the subsidiary maxima will gradually become more prominent, leading to the splitting of the minima into doublets (fig. 5 (c)). As x is decreased β_1 may pass through values $n\pi + \epsilon$ where a fringe reversal takes place. At $x = 0$ the maxima are all equal in intensity, and the maxima and minima are equally spaced at distances $\frac{1}{2} t_0 \cot \phi$.

The general intensity of the fringes, and the contrast relative to the background also change considerably with the value of x. It appears from fig. 4 that in the dynamical region, i.e. for $|x| < 1$, the difference in intensity between the minima and maxima of the fringes can be of the order of unity, i.e. of the intensity of the incident beam, irrespective of

Fig. 5

(a)

(b)

(c)
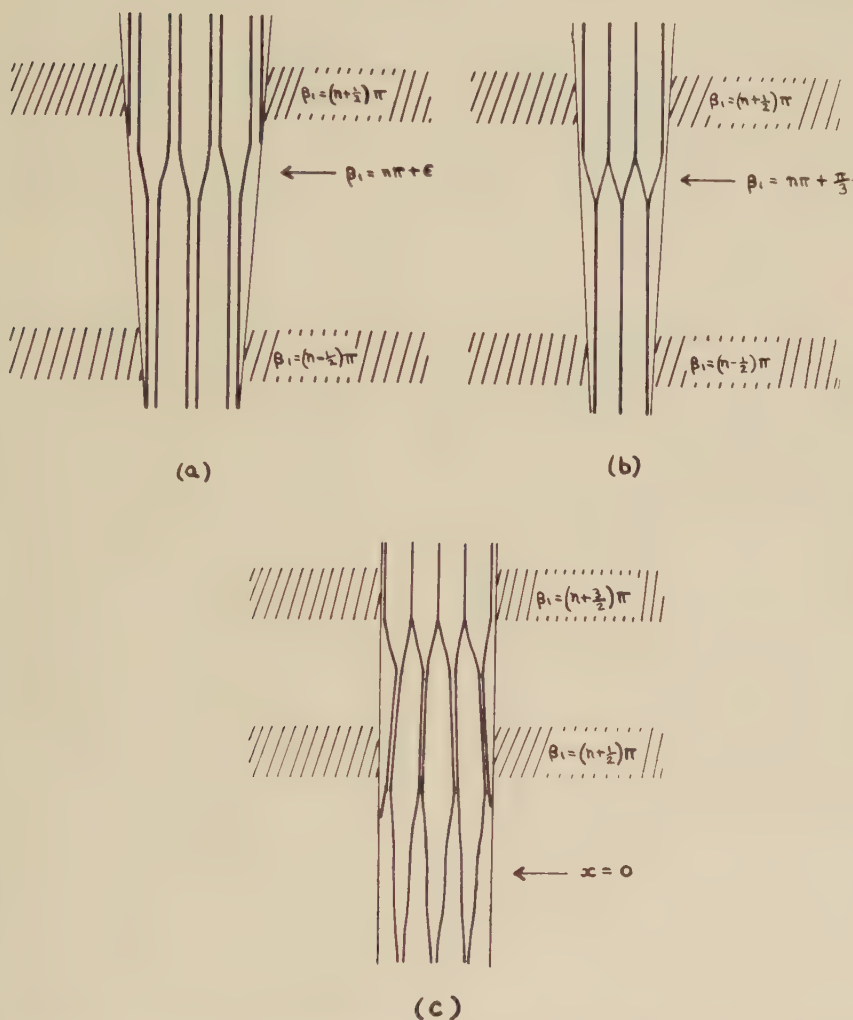
Schematic diagram obtained by a study of fig. 4 showing the change of appearance of the fringes with varying thickness and orientation. The lines represent minima in the intensity profile, and the positions of subsidiary minima of the extinction contours are indicated by shading.

(a) Showing the doublet nature of the fringes in the dynamical region, $|x| < 1$, and the reversal of the fringes at $\beta_1 + n\pi = \epsilon$. The thickness, and therefore β_1 change continuously along the fault.

(b) Appearance of single fringes and their reversal in the kinematical region, $|x| \gg 1$. Near the reversal the fringes are very weak.

(c) Change of appearance of the fringes with orientation, i.e. x varying, for constant thickness. Note the gradual increase in fringe separation as the dynamical region is approached, and also the gradual splitting of the fringes into doublets. At $x=0$, i.e. at the centre of the extinction contour, the fringe separation is halved.

the particular value of the transmitted intensity in the unfaulted crystal. The contrast of the fringes in a region of foil with orientation near a reflecting position is therefore expected to be very great. However, for $|x| > 1$, as x increases towards the kinematical region, the contrast decreases, as can be seen by comparing the curves for $x=1$, and $x=2$ in fig. 4. In the kinematical region, eqn. (26) of paper I shows that the intensity and contrast of the fringes vary as x^2, so that as x increases the fringes should become weaker and eventually become unobservable for sufficiently large x.

§ 4. Experimental Technique

Thin specimens of austenitic stainless steel were prepared by the electropolishing technique described by Bollmann (1956, 1957). Material from two different sources was used, there being no apparent difference between the two samples. The first, supplied by Alleghany Ludlum (Pittsburgh, U.S.A.) in sheet 0·2 mm thick, contained 18 to 20% Cr, 8 to 11% Ni, <0·08% C, <2·0% Mn, <0·04% P, <0·03% S, <1·0% Si and the rest Fe. The second, supplied by Samuel Fox and Co. (Sheffield) in sheet 0·0034 in. thick, was of unknown composition except that it was of the 18/8 type with about 0·10% C. As supplied both these materials were in a well annealed state and contained very few dislocations and stacking faults ; the specimens were therefore rolled up to 5% before thinning. The foils were examined directly in a Siemens and Halske Elmiskop I operating at 80 kv. Both the stacking faults originally present in the worked material, and those produced by the splitting of dislocations into partials during the observations, were photographed on Ilford lantern contrasty plates at instrumental magnifications varying between 20 000 and 40 000. Selected area diffraction patterns were taken using intermediate apertures which isolated areas of diameters 0·25, 0·85 and 2·0 μ on the specimen. Objective apertures of 30 and 50 μ diameter were used so that Bragg reflections did not contribute to the final image. The double condenser lens system with a 400 μ diameter aperture in condenser 2 was used to illuminate the object with a spot of about 5 μ diameter. This produces a divergence of about 3×10^{-3} radians at a point on the specimen at cross-over. Electron micrographs were taken with condenser 2 slightly defocused, so that the divergence of the illuminating beam was less than 3×10^{-3} radians. Since the divergence is then about an order of magnitude less than the Bragg angles involved ($>10^{-2}$ radians) it is a sufficiently good approximation to regard the incident electron beam as a plane wave.

§ 5. Experimental Results

5.1. *General*

Figures 6–11, 13–19 are typical micrographs and diffraction patterns from regions showing the fringes attributed to stacking faults. Selected area diffraction patterns show that the fringes are parallel to traces of

(111) planes on the surface of the specimen.   Since this has been observed for several different orientations it follows that the faults lie on (111) planes, as expected.

Many of the micrographs show dark extinction contours (figs. 6, 9, 11, 13) ; most of these are bend contours, i.e. they are produced when the orientation of the foil is such that electrons are strongly diffracted.   The presence of these contours indicates that the foil is buckled and that the orientation may vary by several degrees over the field of view.   This buckling however is not too severe, as the fringes usually appear as straight lines on the micrographs.

The theory developed in I applies only to the case of one strong reflection.   Diffraction pattern from areas $0\cdot25\,\mu$ in diameter, which usually include only one extinction contour, show that in fact generally only one strong reflection occurs in any particular local region, e.g. fig. 15 (a, d), so that the application of the theory is justified.   Sometimes more than one reflection is present e.g. fig. 15 (b, c) ; nevertheless the theory for one strong reflection generally accounts satisfactorily for the nature of the fringes observed.   Diffraction patterns from larger areas $(\sim 2\,\mu$ in diameter) are usually cross grating patterns consisting of several strong Bragg reflections (e.g. fig. 19), but the individual reflections arise from different areas in the foil.

In general the contrast of the fringe system is particularly good near intense extinction contours (usually low order reflections) in agreement with the dynamical theory.   In regions where extinction contours are not prominent the contrast is weak.

5.2. *Detailed Examination of Fringes*

5.2.1. *Fringes near $x=0$*

Figures 6, 7 and 9 are typical examples of fringes in the dynamical region ($\mid x \mid < 1$).   In fig. 6 two stacking faults occur side by side ; the doublets of the minima (dark fringes) characteristic of the dynamical region are clearly visible (compare fig. 5 (a)).   The contrast is as predicted over the whole of the fault ;   this is typical of faults produced during the deformation prior to thinning.   For faults produced during observation in the microscope by the movement of partial dislocations (fig. 7) the contrast is as predicted in the centre of the bands, but deviations occur at the edges ;   for example in fig. 7 the subsidiary maxima are absent at the edges.   It is thought that these deviations are due to local lattice curvature at the surface produced in the following way. When the partial dislocation moves, a monatomic step would be produced on the surface in the absence of an oxide surface layer.   The oxide layer however may not yield, and as a result a dislocation line parallel to the surface is left at, or just below, the oxide-metal interface.   The strain field of this dislocation will change x locally and change the character of the fringes at the surface.   A similar mechanism has also been used to explain the trails left behind when complete dislocations have moved

(Whelan *et al.* 1957). There is now some experimental information available supporting this view, which will be described elsewhere, but no detailed theoretical analysis has yet been made.

5.2.2. *Variation with thickness*

In regions where the thickness of the specimen varies, but where no bending occurs, it is possible to study the fringe profile as a function of thickness, i.e. of the parameter β_1. In these regions x is therefore constant, but unfortunately it is not possible to determine its value. Figure 8 is a typical example where the main effect is due to thickness variation, which is evident from the general taper of the stacking fault bands at D, E and F. Some of the subsidiary maxima of the intensity in the unfaulted region (i.e. where $\beta_1 = n\pi$) can be seen at A, B, and C. Close to the points where the maxima run into the faults a reversal takes place in a way which agrees generally with the theoretically predicted reversal sketched in fig. 5. Again, however, deviations occur at the edges of the fault, e.g. at C, and these are presumably due to the oxide film effect mentioned above.

5.2.3. *Effect due to buckling of the foil*

In most areas the foil is buckled but of uniform thickness, so that in fig. 4 the appropriate fringe profiles are obtained by considering curves for which both x and β_1, vary, the relation between them being $\beta_1 = \pi t \sqrt{(1+x^2)}/t_0$. In fig. 9, the fringes are typical of the kinematical region at A. On going towards the extinction contour the fringe separation increases, with decreasing x. For small x the subsidiary maxima and the associated doublet structure appears, e.g. at B, as predicted theoretically (fig. 5 (*c*)), until at $x \simeq 0$ (at C), the spacing of the fringes is half that of the composite fringes just outside the central region of the contour. It is generally rather difficult to observe this latter behaviour since the contrast at the centre of most extinction contours in thick regions is usually poor. The features at D, E, F are due to dislocations on nearby planes ; these cause additional strains in the lattice which give rise locally to a modification of the fringe system in the region of overlap. The sharp reversal at G is probably due to overlapping stacking faults and will be discussed later.

Figure 10 illustrates the way in which the number of fringes changes with varying x. At A x is small and the doublet structure is seen. At B x is larger and the subsidiary maxima have disappeared. C corresponds to a value of $\beta_1 = n\pi + \epsilon$ where a fringe reversal takes place. Every time a fringe reversal takes place the number of dark or light fringes changes by one. Another fringe reversal takes place at D, where presumably $\beta_1 = (n+1)\pi + \epsilon$. It follows from the relation between β_1 and x that for large t, x need not change much between neighbouring reversals. In relatively thick regions such as in fig. 9, reversals therefore take place frequently.

Both figs. 9 and 10 show clearly the decrease in contrast of the fringes with increasing x, as predicted by the theory (see § 3).

In thin regions ($t \sim t_0$) only one or two fringes are expected from the curves of fig. 4 : an example of this type is shown in fig. 11. The normal to the foil is [312] and the stacking faults occur on (111). The wedge angle is about 52°, and from the width of the trace the thickness turns out to be about 230 Å. This is about the same as the extinction distance for a (111) reflection (see § 5.3). The theoretical form of the extinction contour is given by putting $t=t_0$ in the expression for the transmitted intensity as a function of x and β_1, i.e. $1 - \sin^2 \{\pi \sqrt{(1+x^2)}\}/(1+x^2)$. This function, shown in fig. 12, predicts the intensity variation of the bend contour in fig. 11 at the centre, i.e. in the dynamical region. The subsidiary maxima of the contour are however poorly defined. This

Fig. 12
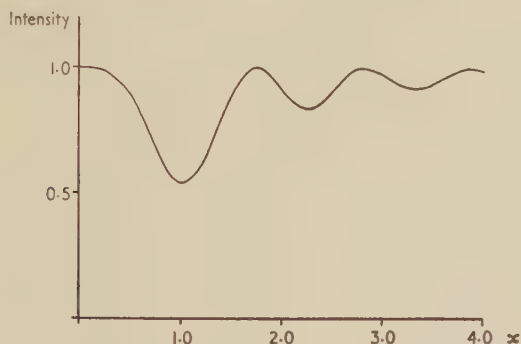
Intensity variation in a bend contour for $t=t_0$. The intensity of the transmitted beam is plotted as a function of x.

may be due to the effect of other weak Bragg reflections in the kinematical region, which are neglected in the simple theory, and to the finite beam divergence the effect of which will be discussed later (§ 5.7). Careful examination of the curves of fig. 4, allowing for the variation of x and β_1 across the contour, shows that the reversals visible at A, B and C occur close to subsidiary maxima of the bend contour in a manner predicted qualitatively by theory. The light fringe at the stacking fault at the centre of the extinction contour near A is a consequence of the doublet character of the fringes near $x=0$, $\beta_1=\pi$, as may be verified by inspection of fig. 4. The complicated behaviour of the fringes at D and other places may possibly be due to the inadequacy of the theory in regions where two or more reflections are operating. It is seen however that in regions where t is about equal to t_0, the contrast at the faults in the dynamical region gives the impression of pairs of parallel light and dark lines. Similar contrast effects have been observed by Menter and Pashley (private communication) on micrographs of evaporated single crystals

of gold. The lines were observed to appear and disappear suddenly in contrast to the dislocation slip traces which faded slowly. On this theory the lines are simply interference effects produced at the stacking fault when partial dislocations separate.

Figure 11 again shows clearly how the contrast of the fringes fades with increasing x, i.e. with increasing distance from the extinction contour.

5.2.4. *Effect of tilting the object*

The variation of the fringe profile with x can also be studied by tilting the object in the microscope. Figure 13 (a), (b), (c) and (d) shows the effect of tilting through a few degrees about the axis shown in fig. 13 (a). The relative positions of the (111) and $(\bar{1}\bar{1}\bar{1})$ reflections, to the axis of tilt are also indicated. These reflections are responsible for the contrast in the sequence. Electron diffraction patterns show in fact that the contrast in fig. 13 (a) is produced by (111), and that in fig. 13 (c) by $(\bar{1}\bar{1}\bar{1})$, and that the most likely explanation of the sequence is that the foil has been tilted from the (111) reflecting position (fig. 13 (a)), through the $(\bar{1}\bar{1}\bar{1})$ reflection position (fig. 13 (c)), to the kinematical region fig. 13 (d). In the thicker regions of which this is an example, the subsidiary maxima of the extinction contours are often not observed, e.g. fig. 13 (a). This is probably due to the finite electron beam divergence, the effect of which will be discussed in § 5.7. The complicated form of the fringe profile in this region at A and B also indicates that x is small. Figure 13 (b) shows that the subsidiary maxima of the fringes disappear as the area is tilted out of the (111) reflecting position. Figure 13 (c) corresponds to a position close to the $(\bar{1}\bar{1}\bar{1})$ reflection. It can be seen that the number of fringes is reduced by one compared with fig. 13 (b), i.e. a reversal of the contrast has most likely occurred, so that the dark fringes of fig. 13 (b) become the light ones in fig. 13 (c) ; the light fringes are clearly visible at the edges of the faults on the original plate, and closer examination shows that they are also present at A and B. Finally in fig. 13 (d) the foil has been tilted almost completely out of the reflecting position ; very closely spaced fringes occur and the contrast is low. The spacing of the fringes in fig. 13 (d) measured on the original plates is about 84 Å, compared with about 168 Å on fig. 13 (a). Assuming that in fig. 13 (a), $x \simeq 0$, and using the relation $t_0' = t_0/\sqrt{(1+x^2)}$, x is found to be about 1·73, a value already within the kinematical region.

5.3. *Fringe Spacing*

The periodicity of the fringes at the faults in the general case $(x \neq 0)$ is given by $t_0' \cot \phi$, where

$$t_0' = t_0/\sqrt{(1+x^2)}, \text{ and } t_0 = (\Delta k_0)^{-1} = |U_g|/k \cos \theta \simeq |U_g|/k = \lambda E/V_g$$

for high energy electrons (see I for notation).

For 80 kv electrons, $\lambda = 0.042$ Å, and V_g is given theoretically by

$$V_g = \{300e\, d_g^2/\pi V\} \sum_j (Z_j - f_j) \exp\{-2\pi i\mathbf{g} \cdot \mathbf{r}_j\}.$$

where e is electronic charge in e.s.u.; d_g is the spacing of the lattice plane, giving the Bragg reflection **g** ; V is the volume of the crystallographic unit cell ; Z_j and f_j are the atomic number and x-ray scattering factor of the jth atom situated at $\mathbf{r}_j$ in the unit cell, and where the summation is taken over all atoms in unit cell.

For stainless steel, assuming a lattice constant of 3·56 Å, and taking the average atomic scattering factor to be the same as that of iron, the following values of lattice potentials and extinction distances are obtained :

V_{111}	V_{200}	V_{220}	$t_{0(111)}$	$t_{0(200)}$	$t_{0(220)}$
14·7	12·7	8·1 volts	228	263	413 Å

A study was therefore made of the fringe spacing in several cases where the orientation was determined sufficiently accurately by electron diffraction. The faults were assumed to lie on (111), in determining the wedge angle ϕ. Plates were chosen which showed prominent low order extinction contours whose indices had been determined by selected area diffraction, and the fringe spacing was measured directly from the micrographs close to the centre of the contours, where x is likely to be zero.

Table 1 is a summary of the results obtained for the {111} and {200} extinction contours. It can be seen that most of the values lie in range 200–250 Å. In particular the value obtained on plate 1127 of 254 Å refers to the point C on fig. 9, where fringes corresponding to half the extinction distance are visible. At this point $x=0$, so that this estimate is likely to be particularly accurate. The low values elsewhere in the table are probably due to deviations from the Bragg position $x=0$. However, since the orientations determined graphically by electron diffraction are not accurate to more than a few degrees, none of the results can be regarded as very accurate. There is therefore no significant difference in the estimates of $t_{0(200)}$ and $t_{0(111)}$. However, the results are in general agreement with the theoretical values above.

5.4. *Reflections with* $\alpha=0$

According to the kinematical and dynamical theories the fringe contrast should vanish in those regions for which the Bragg reflection excited has zero phase factor i.e. $\alpha=0$. Experiments designed to confirm this prediction provide a method of testing for the correctness of the assumed shear vector at the fault, since α depends directly on this quantity. Applying the rules given in § 2 it is found that all of the low order reflection {111} and {200} except for (111) and ($\bar{1}\bar{1}1$), have $\alpha=\pm 2\pi/3$ and should show contrast. Diffraction spots corresponding to the (111) and ($\bar{1}\bar{1}1$) reflections are only produced when the faulting plane is perpendicular to the foil surface, and in this case the dynamical theory as developed cannot be applied. Therefore no positive test is obtained by studying the low order {111} and {200} reflections ; in

practice all these reflections give good contrast at the faults in agreement
with theory.   A positive test may be obtained by studying higher order
reflections such as {220} and {311}, some of which should show contrast

Table 1.   Values of extinction distances, t_0, obtained from fringe spacings
near $x = 0$.   The faulting plane is $(1\bar{1}1)$ in all cases

Plate Number	Reflection	Normal to foil		ϕ degrees	t_0 Å
975	{200}	Near	[201]	39	224
1125	{200}		[201]	38	270
1250	{200}	Near	[101]	42	158, 190, 246
1050	{111}		[312]	52	206, 220, 215, 224
1084 ⎱ 1085 ⎰	{111}	Between [312] and [211]		55	270, 270, 270, 277, 284 270, 277, 298
1095	{111}		[312]	52	208, 210
1105 1107 1109 1111 1112	{111}	Near	[101]	About 45	Average ~200 Maximum value ~230
1127	{111}	Between [312] and [211]		53	254†, 270, 192
1163	{111}	Near	[312]	52	216, 216, 223, 223
1201 ⎱ 1203 ⎰	{111}		[312]	52	170, 175, 223, 223

† This value was obtained at C, fig. 9, where the doublet fringes are visible.

while others should not.   Experiments were therefore performed on
areas of the foil showing many extinction contours and long stacking
faults traversing most of the field of view.   The extinction contours were
indexed by taking electron diffraction photographs from areas of $0 \cdot 25\,\mu$
in diameter straddling the contours as in fig. 14 and fig. 15 (*a*), (*b*), (*c*)
and (*d*).   The results of several cases are given in table 2.   The faulting
plane was derived from the orientation determined by electron diffraction,
and the reflections effective and ineffective in producing contrast on the
micrographs at the faults are worked out for this particular faulting
plane (which is not in general (111) for the orientation stated).   Those

reflections which are underlined should produce no contrast theoretically, i.e. $\alpha = 0$. All others should show contrast. An examination of the experimental observations in table 2 shows that most of the reflections

Table 2. Contrast at extinction contours, as a function of index of the contour. The underlined reflections should show no contrast theoretically

Plate numbers	Orientation normal to foil		Faulting plane	Extinction contours with contrast	Extinction contours without contrast
1188–1200	Near	[211]	(1$\bar{1}$1)	(1$\bar{1}\bar{1}$) and ($\bar{1}$11) (02$\bar{2}$), (0$\bar{2}$2)	(11$\underline{\bar{3}}$), ($\underline{\bar{1}$1$\bar{3}}$)
1142–1146		[211]	(1$\bar{1}$1)	(1$\bar{3}$1), (0$\bar{2}$2)	
1159–1166		[312]	(1$\bar{1}$1)	($\bar{1}$11), (1$\bar{1}\bar{1}$)	
1201–1204		[312]	(1$\bar{1}$1)	($\bar{1}$11), (1$\bar{1}\bar{1}$)	
1171–1187		[101]	(111)	(1$\bar{1}\bar{1}$), ($\bar{1}$11) (0$\bar{2}$0), (020)	$\underline{(\bar{1}31)}$, $\underline{(1\bar{3}1)}$
1211–1220	Between and	[100] [311]	(111) (1$\bar{1}$1)	(0$\bar{2}$0), (020) (0$\bar{2}$0), (020)	($\bar{1}$33) ($\bar{1}$13)
1243–1257	Near	[101]	(1$\bar{1}$1)	($\bar{1}$11), (1$\bar{1}\bar{1}$) (020), (0$\bar{2}$0)	(13$\bar{1}$) $\underline{(1\bar{3}\bar{1}), (\bar{1}31)}$
1229–1240	Between and	[201] [101]	(1$\bar{1}$1)	(020), (11$\bar{1}$)	$\underline{(13\bar{1})}$

lie in the correct category as predicted by theory, with the exception of those on plates 1211–1220. The results for this case are however inconclusive. The results obtained for plates 1188–1200 are particularly interesting and are shown in figs. 16 and 17. In this case two pairs of higher order reflections, i.e. (02$\bar{2}$), (022) and (113), ($\bar{1}\bar{1}$3) are visible. Figure 16 shows that contrast is visible at the stacking faults near the (02$\bar{2}$) and (022$\bar{}$) extinction contours but not near the (1$\bar{1}$3) and (113) contours. This is in agreement with the theory for the type of stacking fault assumed. Figure 17 is a low magnification picture of this area. It is clear that the contrast in the faults is greatest near the (1$\bar{1}\bar{1}$) and (1$\bar{1}$1) extinction contours. It is a general rule that the contrast is greatest for the strongest low order reflections. This point will be discussed in §5.7. The contrast near the weaker (02$\bar{2}$) and (0$\bar{2}$2) contours can be clearly seen. It is very striking how the contrast fades at the (113) and ($\bar{1}\bar{1}$3) contours, even though their order is only slightly higher than

that of the (022) reflection.   The results of this section are therefore in agreement with the predictions of the theory as applied to stacking faults in the face-centred cubic lattice.

5.5. *Diffraction Patterns*

Electron diffraction patterns taken from large areas with considerable numbers of stacking faults as in fig. 18 show spots with streaks running through them.   Figure 19 taken from the area indicated in fig. 18, shows examples of such streaks.   The selected area in this case contained regions on both sides of the grain boundary in fig. 18 so that fig. 19 is a double crystal pattern ; however, it is easy to identify the reflections belonging to each crystal for in general the streaks are found to run almost perpendicular to the stacking fault bands, which are not parallel on the two sides of the grain boundary.   The diffraction patterns are corrected for the azimuthal rotation of the intermediate lens in all cases. The areas A and B on fig. 18 give rise to the A and B spots on fig. 19. The streaks run in directions on the diffraction patterns which are

Fig. 20
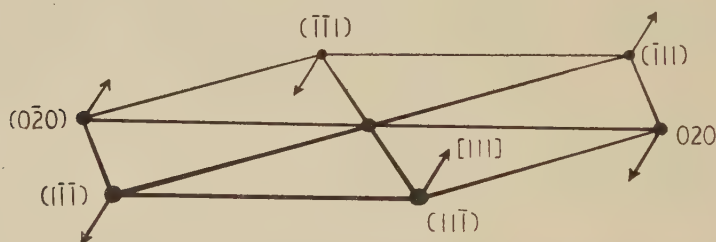
Perspective drawing of (101) plane of the reciprocal lattice of a face-centred cubic structure showing the displacements of the spots in the [111] direction, caused by a distribution of faults on (111).

consistent with stacking faults on (111).   The kinematical theory of diffraction from a random distribution of stacking faults in close packed lattices has been considered by Paterson (1952), who showed that reflections of the form $L=3p$, $(H-K)=3q$ (fig. 3) are unaffected by faults on (111) ;   those for which $L=3p+1$ and $3p-1$, $H-K=3q+1$ and $3q-1$, are broadened and displaced (see § 2).   The displacement is such that the reflections for which $\alpha=+2\pi/3$ move down the [111] axis towards the twin spots while those for which $\alpha=-2\pi/3$ move up the [111] axis. In fig. 18 the orientation of the normal to the foil in A is close to [110] as shown by the cross grating pattern of fig. 19.   Assuming that the faults lie on (111) planes, the directions of the streaks in reciprocal space passing through the points of the (101) plane of the reciprocal lattice are shown in fig. 20.   The arrow indicates the sense of displacement along [111] of the broadened reflections.   In a perfectly flat crystal, electron diffraction patterns would then show single spots where the streaks

intersect the reflecting sphere. However, since the foil is buckled about an axis which lies approximately in the mean plane of the foil, the streaks in reciprocal space are projected normally on to the reflecting sphere and become visible on the electron diffraction photographs. In this way it is possible to explain the streaks through the $(11\bar{1})$ and $(\bar{1}\bar{1}1)$ reflections, fig. 19. The occurrence of a faint streak through the $(20\bar{2})$ reflection, where according to this theory there should be no streak, can be explained qualitatively in terms of secondary Bragg reflection (Cowley *et al.* 1951).

In § 5.4 of I (fig. 7) it was shown that the dynamical theory also leads to streaks in reciprocal space even for a single fault. The main directions along which the interference function has appreciable values in the kinematical regions were respectively the normal $\mathbf{n}_1$ to the foil surface and $\mathbf{n}_2$ to the faulting plane, although a faint streak was also shown to be present along a third asymptote. The first streak along $\mathbf{n}_1$ is mainly due to the finite crystal thickness, while the second along $\mathbf{n}_2$ is due to the stacking fault. When there is more than one stacking fault present the situation is more complicated. The interference functions from the different regions of crystal interfere with each other. However, if the faults all lie on parallel planes only the amplitudes of the streaks are affected, but not their directions. Qualitatively therefore it may be expected that the presence of many stacking faults as in fig. 18 will give rise to a fairly intense streak along the [111] direction in the kinematical region, in agreement with the Paterson theory.

Electron diffraction photographs of very small areas, over which buckling is not appreciable, should therefore show double spots in the kinematical region, corresponding to the intersections of the streaks along $\mathbf{n}_1$ and $\mathbf{n}_2$ with the reflecting sphere.

This effect is shown in fig. 15 (*a*), (*b*), (*c*), (*d*) which were taken from the $0.25\,\mu$ diameter areas shown in fig. 14. The diffraction photographs were taken of areas covering different extinction contours, so that different strong Bragg reflections occur in each case. The splitting of the various weak diffraction spots into doublets can be clearly seen, e.g. for the (111) reflections at A, B, in fig. 15 (*a*). The spots closest to the (111) diffraction ring at A and B correspond to the spike in reciprocal space along $\mathbf{n}_1$, while those furthest away are produced by the spike along the normal $\mathbf{n}_2$, to the fault planes. The separation of the spots in the doublet in fig. 15 (*a*) at A corresponds to an interference fringe spacing of 21 Å in the final image. The intensity and spacing however are so low that the fringes are not resolved on the micrographs. The mean orientation of the normal to the foil in fig. 14 is [101] and the faults lie on (111) or $(1\bar{1}1)$ so that the wedge angle ϕ is about 35°. The spacing of the spots in the doublet at A therefore corresponds to a distance s of 0.07 Å^{-1} (fig. 7 of I). Taking the lattice constant to be 3.56 Å this compares with the value 0.49 Å^{-1} for the reciprocal distance d^*_{111}. Using the relation $x = t_0 s$ and taking $t_{0(111)} \sim 230$ Å, it is found that $x \sim 16$ which

is outside the dynamical region ($|x| <1$). The value $x=1$ for the (111) reflection in stainless steel corresponds to a deviation of $0\cdot5°$ from the Bragg angle. It follows that the deviation from the Bragg angle for the doublet at A is about $8°$ which is considerable.

The difference in spacing between the doublets at A and B, fig. 15 (*a*), can be explained quantitatively in terms of the curvature of the reflecting sphere. The ratio obtained for the spacings at A and B is $1\cdot20$. In fig. 21 the curvature of the reflecting sphere is taken into account and it follows that $s'=d^{*2}_{(111)}/2\chi=0\cdot005$ Å^{-1}. Then to a good approximation the ratio of the spacing should be $s/s-2s'=0\cdot07/0\cdot06=1\cdot17$ which is sufficiently close to $1\cdot20$ considering the error in measuring the doublets.

Fig. 21
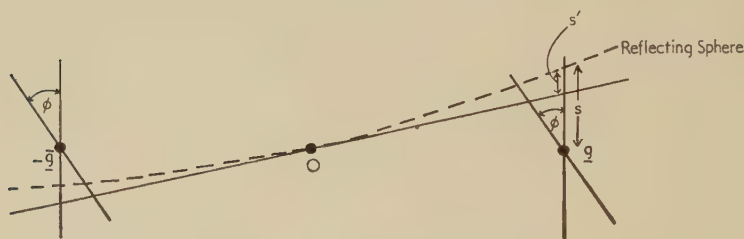
The reflecting sphere construction ; streaks in reciprocal space give rise to double spots, and the spacing between the components of the doublets is different on the two sides of the origin of the reciprocal lattice owing to the curvature of the reflecting sphere.

Near the Bragg reflecting position there should be three closely spaced spots corresponding to the spikes A_1 A_2 A_3 (fig. 7 of I). However it is not possible to resolve these spots on the diffraction photographs owing to the finite beam divergence.

5.6. *Overlapping Stacking Faults*

It was shown in I, § 6.3, that the case of two stacking faults on neighbouring planes could be treated by adding the vector displacements at each fault and by considering the shear to take place on one plane. This treatment then leads to a combined phase factor which is equal to the sum of those produced by the individual faults. Consider a reflection for a face-centred cubic crystal for which the phase factor for an intrinsic fault is $+2\pi/3$. Then, if a second intrinsic fault is placed on a neighbouring plane to the first, an extrinsic fault is formed. The total phase factor will then be $2\pi/3+2\pi/3=4\pi/3$. This is also equivalent to a reversal in sign of the original phase factor, i.e. $+2\pi/3$ changes to $-2\pi/3$. This treatment also holds for the case of two intrinsic faults on planes which are close together but not neighbouring, so that the combined fault need not necessarily be extrinsic. Similarly, in a face-centred cubic lattice, three overlapping faults produce zero total phase factor.

For the case of two overlapping faults, examination of eqn. (24) of the previous paper, and of eqn. (1) of this paper shows that a reversal in sign of α is equivalent to a change of sign of x, so that in fig. 4 the dotted curves become the full curves and *vice versa*. It can then be seen by inspection of fig. 4 that reversal of the intensity profile then occurs when $n\pi + \epsilon < \beta_1 < (n+1)\pi - \epsilon$, in the sense that strong maxima of the fringe profile become weak maxima and *vice versa*, and that when $n\pi - \epsilon < \beta_1 < n\pi + \epsilon$ the intensity may alter but reversal in the above sense does not occur. In the kinematical limit when x is large ϵ tends to $\pi/3$ so that complete reversal occurs when $n\pi + \pi/3 < \beta_1 < n\pi + 2\pi/3$ otherwise only the intensity changes. The reversal is complete in this case in the sense that the weak maxima of the dynamical region are not present. The reversal when two stacking faults overlap occurs abruptly at the dislocation line where the faults begin to overlap, in contrast to the gradual type of reversal which occurs in the neighbourhood of the subsidiary maxima of the extinction contours, and can easily be distinguished from it. Clear examples of reversals due to overlapping faults can be seen in fig. 22 (*a*) and (*b*). In these cases, $n\pi + \epsilon < \beta_1 < (n+1)\pi - \epsilon$ since the reversal is a complete one. Very clear cases of overlapping stacking faults can also be seen on fig. 3 of Bollmann (1957).

In fig. 10 at A an example occurs of an overlapping stacking fault in the region $n\pi - \epsilon < \beta_1 < n\pi + \epsilon$. It can be clearly seen that the fringe profile changes but does not reverse. Subsidiary maxima are visible in one region which are not visible at all in the other.

Areas showing no contrast, such as at B in fig. 22 (*a*) and (*b*) are probably produced by the overlap of three stacking faults. Figure 23 shows the effect of overlap of a number of stacking faults. The interpretation is as follows. At A the fringes are produced by one stacking fault ; at B the fringes are reversed due to overlap of two faults ; at C the contrast vanishes due to overlap of three faults, or possibly due to absence of faults ; this sequence then repeats a number of times. Energy considerations and the nature of the array suggest that it consists of a number of overlapping stacking faults, the number starting from one at A and increasing by one at every change in the fringe system. The number of overlapping stacking faults is marked on fig. 23. For $3n+1$ overlapping faults (n is an integer) α is effectively $2\pi/3$; for $3n+2$ faults, $\alpha = -2\pi/3$, and for $3n$ faults, $\alpha = 0$, so that as n increases the sequence observed on fig. 23 is produced. The importance of this array in relation to the theory of twinning or of phase transformation will be discussed elsewhere.

5.7. *Effect of Divergence of Beam*

As stated in § 4 the divergence of the incident beam is about 3×10^{-3} radians under operating conditions. The beam therefore consists of a whole set of parallel beams incident at slightly different angles. Each

component beam produces a diffraction pattern corresponding to a particular value of x. The complete diffraction pattern therefore consists of a superposition of patterns of slightly different x. As a result some of the fine detail of the pattern calculated for a parallel beam will be lost. Similarly, the intensity distribution on the micrographs is given by the superposition of a whole range of intensity distributions corresponding to slightly different values of x. Again some of the details of the intensity variations will therefore be averaged out.

The range of values of x corresponding to the divergence of the incident beam is given by $dx = t_0\, ds = (t_0/d)d\theta$ where d is the spacing of the reflection operating. Now t_0 increases and d decreases with increasing order of reflection; dx therefore increases with increasing order of reflection. This accounts for the fact that subsidiary extinction contours are not observable for higher order reflections, and also for the decrease in contrast of the stacking fault fringes with increasing order of reflection (see § 5.4).

This discussion can be made quantitative as follows. For a low order reflection the order of magnitude of dx is $230/2 \cdot 06\, d\theta$ and taking $d\theta = 3 \times 10^{-3}$ radians, $dx = 0 \cdot 34$ which is quite appreciable. Now the periodicity of subsidiary extinction contours is given by $\varDelta\beta_1 = \pi$ (this can be derived from the formulae given in § 6.1 of I); the corresponding range $\varDelta x \sim t_0 \sqrt{(1+x^2)}/(tx)$. Hence $dx/\varDelta x \sim tx\, d\theta/\{d\sqrt{(1+x^2)}\}$. If this ratio $\geqslant 1$, the subsidiary contours will be averaged out. In practice therefore these contours will be less prominent the greater the thickness. In the kinematical region, $x > 1$, the contours will be particularly affected. For an average $t \sim 500$ Å, and a high order reflection $d \sim 1 \cdot 5$ Å, $dx/\varDelta x \sim 1$ in the kinematical region, and it is therefore not surprising that subsidiary maxima are not observed. Even for low order reflections the subsidiary contours, when $x > 1$, should be considerably weaker, and they are in fact hardly observable, particularly in thick regions (see § 5.2.3). The strong central, and in favourable cases also, the adjacent contours in the dynamical region will however be observable, partly because x is small, and partly because of their great intensity. These considerations then account largely for the nature of the extinction contours which are observed.

With regard to the stacking fault fringes, their spacing is given by $t_0 \cot \phi / \sqrt{(1+x^2)}$, and the spread in x is accordingly not very important. However, the fringes reverse in intensity at $\beta_1 = n\pi + \epsilon$ and reversal therefore takes place in a range $\varDelta\beta_1 \sim \pi$, equal to that corresponding to the periodicity of the extinction contours. From the above discussion on the visibility of extinction contours it follows therefore that the contrast of the stacking fault fringes will tend to be poor in the kinematical region, will be best in the dynamical region, and will become weaker with increasing order of reflection. This is in agreement with the observations (see § 5.4). It should be noted, however, that although the weak subsidiary extinction contours may become hardly visible when $dx/\varDelta x \geqslant 1$,

the stacking fault fringes will still be observable, although with less contrast, owing to their comparatively greater intensity.

§ 6. Conclusions

The agreement between the theory of I and the experimental observations described in this paper is satisfactory and suggests that the theory is essentially correct and confirms that the interference fringes observed in stainless steel are due to stacking faults on (111) planes.

With regard to the theory, it is in principle possible to apply a similar treatment to other types of overlapping crystals, the essential step consisting of the matching of the wave functions at the crystal boundaries. In this way it is possible to explain similar types of fringes observed at grain boundaries, and care must be taken in such cases in distinguishing between interference fringes and dislocation lines at the boundaries. Again, it seems possible, at least formally, to develop a dynamical theory for crystals containing imperfections other than stacking faults, by introducing a continuously varying phase factor α, and by considering the crystal as a series of thin slabs. A serious defect of the theory is the neglect of absorption. Formally this can be introduced by postulating complex Fourier potentials, but at present the nature of the complex scattering factor is not known.

On the experimental side, it is clear that the contrast due to single stacking faults is characteristic and very pronounced, so that such faults can be detected by transmission electron microscopy. It should be possible for example to determine whether diffusionless transformations proceed by the movement of partial dislocations.

Since the fringes due to stacking faults are somewhat similar to thickness fringes in wedge crystals, it is necessary to consider the fringes in some detail before they can be attributed to faults. Thin precipitates for example might give rise to somewhat similar fringe patterns (see for example Fukano and Ogawa 1956).

REFERENCES

BOLLMANN, W., 1956, *Phys. Rev.*, **103,** 1588 ; 1957, *Proceedings of the Stockholm Conference on Electron Microscopy* (Stockholm : Almquist and Wiksell), p. 316.

COWLEY, J. M., REES, A. L. G., and SPINK, J. A., 1951, *Proc. phys. Soc. Lond.* A, **64,** 609.

FRANK, F. C., 1951, *Phil. Mag.*, **42,** 809.

FUKANO, Y., and OGAWA, S., 1956, *Acta Cryst.*, **9,** 971.

PATERSON, M. S., 1952, *J. appl. Phys.*, **23,** 805.

WHELAN, M. J., and HIRSCH, P. B., 1957, *Phil. Mag.*, **2,** 1121.

WHELAN, M. J., HIRSCH, P. B., HORNE, R. W., and BOLLMANN, W., 1957, *Proc. roy. Soc.* A, **240,** 524.

Delayed γ-Ray Angular Correlations in Tantalum 181†

By R. E. Azuma and G. M. Lewis

Department of Natural Philosophy, The University, Glasgow

[Received July 31, 1957]

Abstract

The anisotropy of the 133–480 kev γ-ray cascade in ^{181}Ta has been investigated for liquid sources, using a time-sorting coincidence unit of short resolving time. The maximum anisotropy has been found to be $-(39\pm3)\%$, consistent with the $\frac{1}{2}$(E2) 5/2 (E2+M1) 7/2 assignment; and evidence has been obtained for attenuation of the correlation in liquid sources.

§ 1. Introduction

The element ^{181}Ta is of particular interest as it is one of the few elements which possesses a metastable state in the 10^{-6}–10^{-9} sec region, where the effects of nuclear re-orientation under extranuclear forces may be observed. Moreover, the unusual γ-ray transition probability of this state in ^{181}Ta, and the existence of a higher 22 μsec state, make the spin assignments, and determination of the multipolarity of the radiations, of importance from a nuclear model viewpoint.

Angular correlation measurements were made by McGowan (1954) on the dominant 133–480 kev γ-ray cascade of ^{181}Ta, which has an intermediate state of half-life 10^{-8} sec. A maximum anisotropy, $(W180°-W90°)/W90°$, of -38% was observed with a liquid source, with a coincidence unit of resolving time comparable to the half-life. In the interpretation of the results considerable attenuation of the anisotropy was assumed to be present, due to the perturbation of the intermediate metastable state by extranuclear forces, which would modify the experimentally observed Legendre Polynomial coefficients A_2, A_4, of -0.28 and -0.06 respectively, of the correlation function. This led to a spin assignment of 5/2(E2) 9/2(E2+M1) 7/2. The ground state was known to be 7/2 (Gisolf and Zeeman 1933). Heer *et al.* (1955) have briefly reported a similar anisotropy and they suggested a fit with the spin sequence $\frac{1}{2}$(E2) 5/2 (E2+M1) 7/2. Reference should also be made to the work of Paul and Steffen (1955).

The true angular correlation occurs only when there is no delay between the formation of the intermediate state nucleus and the emission of the

subsequent γ-ray. To delimit the spin sequence it seemed therefore desirable to study the anisotropy in this cascade, for any attenuation effects as a function of emission time delays, with an apparatus of very short resolving time. To obtain a reliable determination it would be advantageous to record the γ–γ coincidence rate, at a given angular separation of the detecting counters, for the various emission time delays, all in one period of measurement. A time-sorting method (cf. Neilson and James 1955) appeared to provide promise in this direction, and a system has been developed here which enables γ–γ coincidence data to be acquired at a specific angle setting in an automatic manner.

Furthermore, information would be expected from these experiments concerning the attenuation and mode of perturbation of the intermediate state nucleus under extranuclear fields. With the passage of time the specific spin orientation of the intermediate state nucleus is affected. Abragam and Pound (1953) have shown that the angular correlation $W_\theta(t)$ can be expressed as a function of time t by :

$$W_\theta(t) = \sum_k G_k(t) \cdot A_k P_k(\cos\theta) \quad \ldots \quad \ldots \quad (1)$$

where A_k are the coefficients specifying the angular correlation at angle θ, and $G_k(t)$ are the attenuation coefficients. Theoretical values for these attenuation coefficients have been determined by these authors for static electric quadrupole interactions in polycrystalline powders, and for time dependent interactions in liquids. In the case of solids a periodic variation of anisotropy, with respect to time, is predicted, equivalent to a memory effect, and a hard core value remains when the correlation is integrated over a time large compared with the precession time of the quadrupole moment in the crystalline field gradients. Liquids on the other hand would be expected to show an exponential law of attenuation, resulting in an irrecoverable loss of anisotropy, viz. :

$$G_k(t) = \exp(-\lambda_k t) \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (2)$$

where λ_0 is zero and λ_k $(k \neq 0)$ are specified products of the electrical and relaxation constants of the solution. The γ–γ coincidence rate, $C_\theta(t)$, from a liquid source should therefore obey the relation :

$$C_\theta(t) = \alpha \exp(-t/T)[A_0 + \sum_k A_k \exp(-\lambda_k t) P_k(\cos\theta)] \quad \ldots \quad (3)$$

where T is the decay time of the intermediate state nucleus, and α depends on the source strength. When λ_k is small compared with $1/T$ it can be shown that the plot of $\log C_\theta(t)$ should not differ markedly from a linear plot over a few half-lives, but that the mean slope should differ from $1/T$, and different slopes should occur at different angles.

Finally, fig. 1 illustrates the decay scheme for ^{181}Ta recently proposed by Boehm and Marmier (1956) on the basis of crystal diffraction experiments. The intensity of each of the two weak γ-rays (137 kev and 476 kev), leading into the 480 kev level is about 4% of the intensity of the strong 133 kev γ-ray considered here. Reference too should be made to the recent work on ^{181}Hf by Gimmi *et al.* (1956), Stelson and McGowan (1957), and Snyder and Frankel (1957).

Fig. 1
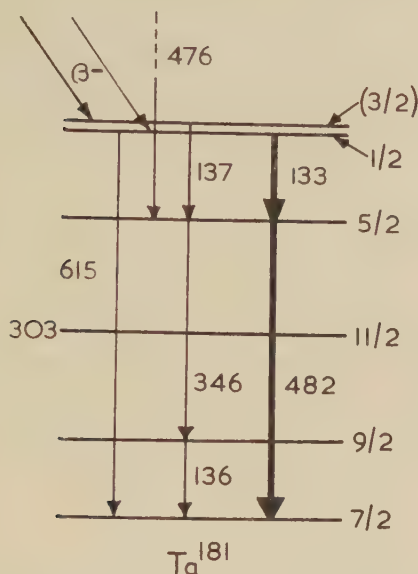
Decay scheme of ^{181}Ta, formed by β-decay of ^{181}Hf.

§ 2. Apparatus

The high speed coincidence circuitry was designed to convert the interval between the arrival of two signals in scintillation counters into a pulse height, the pulse height being a linear measure of the time interval. This is achieved by leading the output from each photomultiplier collector into a limiter valve. The limiter pulses are passed through a variable length of cable, and clipped. The two clipped pulses are amplified and applied to the grids of a gated beam 6BN6 valve. The output of the 6BN6 valve is integrated to give a pulse whose height is proportional to the amount of overlap of the two input pulses. The output pulse height is displayed on a Hutchison–Scarrot type multichannel kicksorter. Figure 2 illustrates the arrangement used in all the work described below. R.C.A. photomultipliers were employed, and E180F valves were used in the limiter stage. 95 ohm coaxial cabling led to the 95 ohm clipping line, attached to the grid of the EFP60 secondary emitter valve, which amplified without change of phase.

It may be mentioned in passing that a development of this system has been made to operate satisfactorily, where amplification by the EFP60 valve is dispensed with. The necessary pulse height to operate the 6BN6 valve is obtained by the use of two E180F valves in parallel in the limiting stage, all the cabling being 200 ohm coaxial line.

In addition to the fast time-sorting coincidence unit, signals from the dynodes of the photomultipliers operated a slow coincidence gating system. In this way only those fast coincidence pulses associated with signals in given energy ranges were recorded (cf. Azuma and Lewis 1955).

The circuit was developed in its initial stages with R.C.A. 6342 photomultipliers operated at 2000 volts, and utilized cylindrical ($1\frac{3}{4}$ in. diam. $\times$ 2 in.) plastic scintillators. Annihilation radiation from ^{22}Na was used as a source of 'prompt' γ-rays, and the side channels were set to accept the high energy portion of the Compton distribution. This showed that resolving times (half-width at half-height), of less than $\frac{1}{2}$ mμsec could

Fig. 2
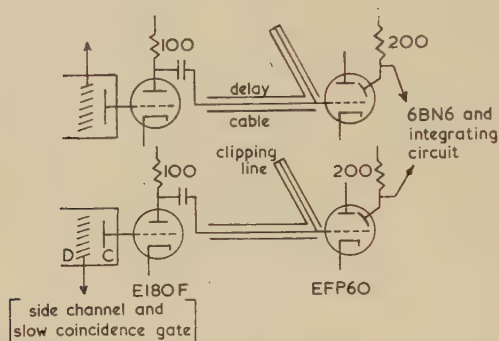
Schematic diagram of counters and input units to fast time-sorting coincidence circuitry.

Fig. 3
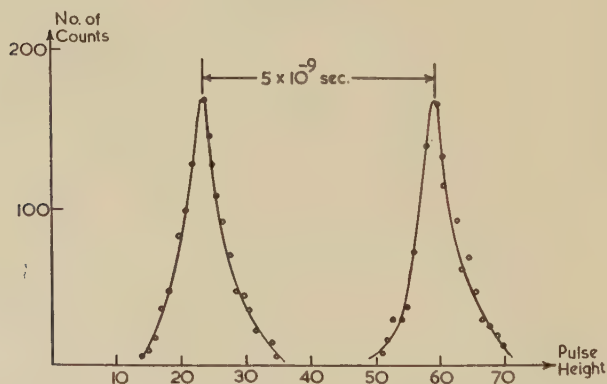
Coincidence resolving curves, as displayed on a multi-channel kicksorter, and obtained with an artificial delay of 5 mμsec between them, with annihilation radiation and plastic scintillators.

be achieved. Figure 3 shows two curves as displayed on the kicksorter in succession. To obtain the second of these curves additional cabling, giving an artificial delay of 5 mμsec, was inserted.

When two sets of coincident signals are fed into the time sorter, complete overlap occurs and the maximum pulse height is produced. If one set of pulses is delayed by a fixed amount with respect to the other set, a fixed overlap would ideally occur giving rise to a constant output

pulse proportional to the overlap. Due to the variation in the time of producing such pulses the output will not be quite constant in height and will generally be symmetric about the mean, except where total overlap occurs. In all the experiments described here the positions approaching complete overlap were always avoided.

§ 3. EXPERIMENTAL PROCEDURE

In the specific study of the radiations of ^{181}Ta it was necessary to use a NaI (Tl) scintillator ($1\frac{3}{4}$ in. diam. × 2 in.) on one side in order to select only the 480 kev γ-ray, with a discriminator in the side channel. A plastic scintillator ($1\frac{3}{4}$ in. diam. × 2 in.) was used on the other side to detect the 133 kev γ-rays. A single channel kicksorter was used to accept a narrow portion of the 133 kev γ-ray Compton distribution. With these settings, and with the use of a thin lead shield in front of the 480 kev γ-ray counter, coincidences due to back scattering were eliminated. Also, all this enabled true prompt coincidence comparisons to be

Fig. 4
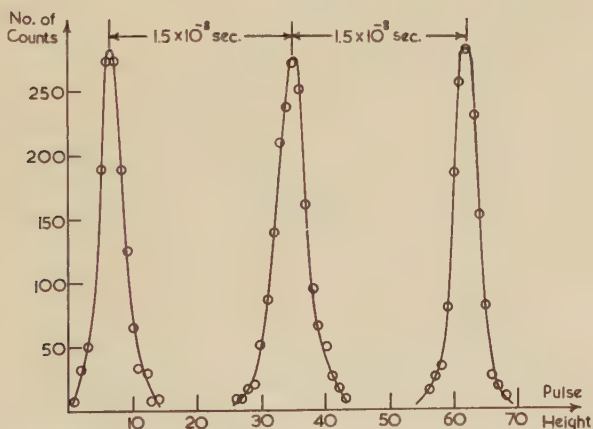
Coincidence resolving curves, as displayed on a multi-channel kicksorter, obtained with an artificial delay of 15 mμsec between each pair, with annihilation radiation as a prompt source. NaI (Tl) was used on one counter, plastic on the other, and side channel selection used so that energies accepted due to the prompt source corresponded to those of the 480 and 133 kev γ-rays of ^{181}Ta.

made with annihilation radiation from a ^{22}Na source. Figure 4 shows three resolution curves obtained under these conditions with annihilation radiation, the time shift between each pair of curves is 15 mμsec. As expected, the curves have the same shape, and the output pulse is substantially proportional to overlap. These features are essential to the method. An overall check on the method could be provided by placing a solid source of HfO_2 sandwiched closely between the two scintillators, so that a large solid angle was subtended. A linear logarithmic plot of counts against delay time could be obtained, with the correct decay time.

In the later experiments, an R.C.A. 6810 tube was used on the 133 kev side to give more gain, which gave similar results to those described above with the R.C.A. 6342 tube. The resolving time (half-width at half-height) was 2 mμsec.

The counters were mounted on a light plywood table. The source was situated 3 in. from the face of each scintillator. Delayed coincidence curves were displayed on the multichannel kicksorter for the counters at 90°, and then at 180°. Total single counting rates were taken as checks. Random coincidence distributions were obtained by placing a long delay in the 480 kev side, and the ratio of random to real counts did not exceed 1/20, and in most runs was considerably less. The anisotropy could then be determined from a comparison of the displays. In all the experiments the sources were contained in thin walled glass or polythene containers of no more than $\frac{1}{8}$ in. diam. The sources of Hf were obtained from A.E.R.E. in oxide form, after irradiation for 1 week in the Harwell pile at pile factors of order 12.
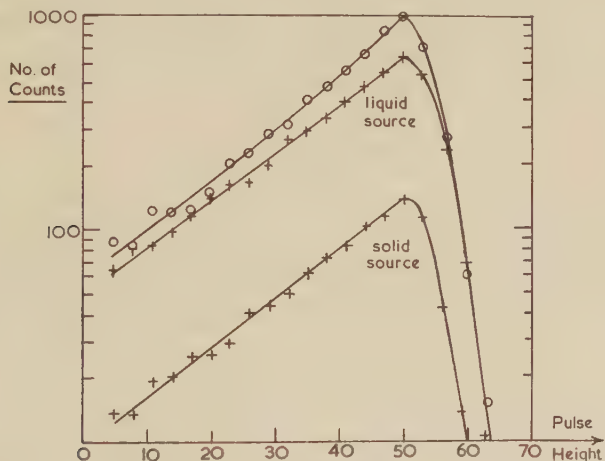
Fig. 5
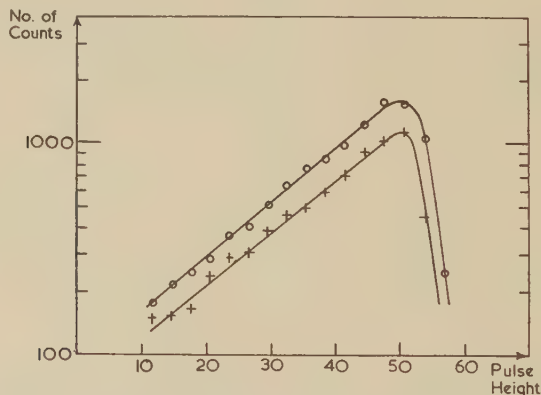
Delayed coincidence curves for a source of HfO_2 dissolved in 27N HF. The upper curve was obtained with the counters at 90°, the lower with the counters at 180°. The slope of the curves correspond to a half-life of 10^{-8} sec.

§ 4. Results and Discussion

A source of HfO_2 was dissolved in 27N HF. A portion of this liquid containing about 1/10 mg of ^{181}Hf was placed in a polythene holder, and two of the curves taken with this source are shown in fig. 5, where the random counting rate has been subtracted. All the curves were obtained within 24 hours of preparing the source, and no deterioration of anisotropy was noticed during this period. The maximum anisotropy, after correction for the finite angle subtended by the source, was $-(39\pm3)\%$. The slopes of the two curves are substantially the same, but evidence for a slight attenuation in this source does exist. The curves extend over a time of three half-lives.

An alternative liquid source was also used which was prepared by fusing HfO_2 with KHF in a platinum crucible, and the product was dissolved in water. Figure 6 shows results with this liquid source, for pairs of runs in succession. The random counts, which were very low, have been subtracted. The observed anisotropy, before correcting for solid angle, was a maximum of -33% decreasing to half of this over a period of some $3\frac{1}{2}$ half-lives of the source. The corrected maximum anisotropy at t zero is -39%. In order to show that the difference between these two curves was not instrumental, plots were obtained with a solid source of HfO_2 in the position previously occupied by the liquid source with the counters at $90°$, and $180°$. This pair of plots were accurately parallel (cf. below), showing that the attenuation in the liquid was a real effect. One of the plots for the solid source is shown in fig. 6, along with the two curves for the liquid source.

Fig. 6

The upper pair of curves are delayed coincidence curves for a source of HfO_2, fused with KHF, and dissolved in water. The top curve of this pair was obtained with the counters at $90°$, the one under it with the counters at $180°$. The lowermost curve is a delayed coincidence curve for a solid powder source of HfO_2 in the position previously occupied by the liquid source. For this curve the number of counts is ten times that of the scale shown.

In all the experiments performed with liquid sources the maximum anisotropy, i.e. at time t zero, was found to be $-(39\pm3)\%$, confirming the spin and transition assignment $\frac{1}{2}(E2) 5/2 (E2+M1) 7/2$, where the second mixed transition is predominantly quadrupole (cf. Frauenfelder 1955). The weak γ-ray at 137 kev will only add to the correlation through its 4% intensity, and the weak 476 kev γ-ray will have a negligible effect because of the gating system used, so that the E2 contribution will lie close to the figure of $97\frac{1}{2}\%$ proposed by Heer *et al.* (1955). On the basis of states satisfying the single particle model, this ratio of

E2/M1 intensities is abnormally high, since in an (E2+M1) transition of this type the M1 term should dominate (cf. Blatt and Weisskopf 1952). The lifetime on such a model would accord more with a pure E2 transition, but a spin assignment of type 15/2(E2) 11/2 (E2) 7/2, for instance, does not give, at all, the angular anisotropy found here. The abnormality seems to arise because the ground state is a rotational state and the 480 kev state is a particle excited state (cf. Alaga *et al.* 1955 also).

Several anisotropy measurements were made on solid HfO_2 sources to determine whether any periodic variation of the anisotropy could be observed. No definite departure could be observed from a linear logarithmic plot, and the overall plots at 90° and 180° were parallel lines, within the limits of measurement. It would appear that the minimum resolving time, $(2 \times 10^{-9}$ sec), attainable was not sufficiently short to display the precessional period T_0 of the quadrupole interaction in the crystalline field gradients, which is of order $\hbar/(eQ \, . \, | \, \text{grad } \mathbf{E} \, |)$.

In conclusion, the relative slopes of the three lines in fig. 6 can be shown to afford good quantitative agreement with the exponential law of attenuation (2), with an appropriate choice of the constant λ_2, $(\lambda_4 = 1 \cdot 7 \, \lambda_2$ for an intermediate state of spin 5/2—cf. Abragam and Pound 1953).

References

ABRAGAM, A., and POUND, R. V., 1953, *Phys. Rev.*, **92,** 943.
ALAGA, G., ALDER, K., BOHR, A., and MOTTELSON, B. R., 1955, *Dan. Mat. Fys. Medd.*, **29,** 9.
AZUMA, R. E., and LEWIS, G. M., 1955, *Phil. Mag.*, **46,** 1034.
BLATT, J. M., and WEISSKOPF. V. F., 1952, *Theoretical Nuclear Physics* (New York : Wiley), p. 627.
BOEHM, F., and MARMIER, P., 1956, *Phys. Rev.*, **103,** 342.
FRAUENFELDER, H., 1955, *β and γ Ray Spectroscopy (Siegbahn)* (Amsterdam : N. Holland Publishing Co.), p. 531.
GIMMI, F., HEER, E., and SCHERRER, P., 1956, *Helv. phys. Acta*, **29,** 147.
GISOLF, J. H., and ZEEMAN, P., 1933, *Nature, Lond.*, **132,** 566.
HEER, E., RUETSCHI, R., GIMMI, F., and KUNDIG, W., 1955, *Helv. phys. Acta*, **28,** 336A.
McGOWAN, F. K., 1954, *Phys. Rev.*, **93,** 471.
NEILSON, G. C., and JAMES, D. B., 1955, *Rev. sci. Instrum.*, **26,** 1018.
PAUL, H., and STEFFEN, R. M., 1955, *Phys. Rev.*, **98,** 231A.
SNYDER, E. S., and FRANKEL, S., 1957, *Phys. Rev.*, **106,** 755.
STELSON, P. H., and McGOWAN, F. K., 1957, *Phys. Rev.*, **105,** 1346.

The Remanent Magnetization
of Single Domain Ferromagnetic Particles†

By E. P. WOHLFARTH and D. G. TONGE

Department of Mathematics, Imperial College, London

SUMMARY

The remanent magnetization is calculated for particles with a number of equivalent easy directions of magnetization, and also for particles with mixed uniaxial anisotropies. Application is made to a variety of materials. In an Appendix the magnetization curves for particles with unidirectional anisotropy are calculated.

---

§ 1. INTRODUCTION

IN certain technological applications of hard magnetic materials, some of which are composed, at least in part, of single domain ferromagnetic particles (for a review see Wohlfarth 1956), it is desirable for the remanent magnetization I_r to be as large as possible. Thus for permanent magnet alloys and powder magnets a large remanent magnetization, if combined with a large coercive force and a relatively square demagnetization curve, leads to a large maximum energy product. Again, for the iron oxide powders composing the coatings of magnetic sound recording tapes, the audio signal carried by the tape after its passage through the recording head is partially determined by the remanent magnetization of the coating. The maximum value of the reduced remanent magnetization,

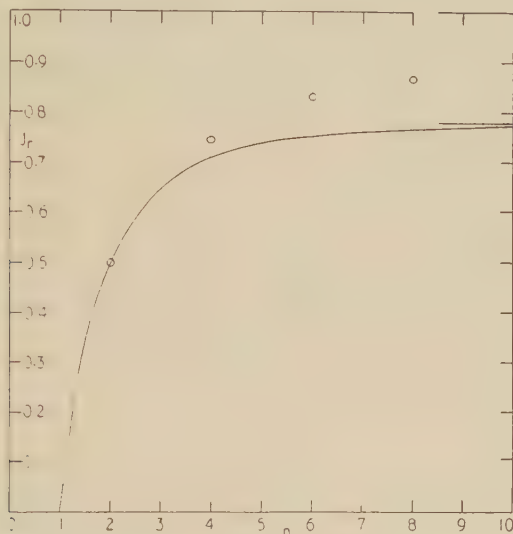$$j_r = I_r/I_0, \qquad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (1.1)$$

where I_0 is the saturation magnetization, is clearly 1. For an assembly of single domain particles with uniaxial anisotropy (due to shape, magnetocrystalline effects, etc.) $j_r = \frac{1}{2}$ if the distribution of the easy axes of magnetization is random. For some of the materials referred to the particles are, in fact, approximately uniaxial, and an increase of j_r is brought about by a magnetic treatment during manufacture. This tends to align the easy axes parallel to the magnetic field applied during the treatment. Complete alignment would raise j_r to 1, but for technical reasons this is not possible, and j_r values about 0·8–0·9 are usually regarded as quite acceptable.

Owing to the importance of the remanent magnetization in the techno-logical applications of single domain materials a more general investigation than previously has been carried out, and the results are reported in this

---

† Communicated by the Authors.

paper.   Earlier investigations on uniaxial particles dealt with the varia-
tion of j_r with the degree of alignment θ of a particle relative to the field
direction, giving (Stoner and Wohlfarth 1948)

$$j_r = \cos \theta, \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1.2)$$

and with the mean value for a random assembly of such particles.   The
variation of j_r as a function of the previously applied field strength was
also discussed (Wohlfarth 1955 a), and the results applied to the coatings
of sound recording tapes.   Here calculations of j_r are presented for
particles having a variety of more complicated types of anisotropy than
uniaxial.   In § 2 a discussion is given for particles with a number of
equivalent easy directions distributed regularly either in a plane or in
space.   In § 3 the effect on the remanent magnetization of the simulta-
neous action of two unequal uniaxial anisotropies is discussed.   This
state of affairs may arise if, for example, both shape and magnetocrystal-
line effects are important.   In an Appendix the magnetic properties
of particles with unidirectional anisotropy are discussed, with reference
to the experiments of Meiklejohn and Bean (1957).

§ 2. Regular Anisotropies

For some applications it is necessary to consider the magnetic properties
of single domain particles for which the anisotropy is not uniaxial.
While it is difficult in such cases to calculate the complete magnetization
curve and, in particular, the coercive force (Néel 1947), the calculation
of the remanent magnetization is quite simple.   In this section results
are given for particles with regularly distributed equivalent easy direc-
tions of magnetization.

2.1. *Planar Distribution of Easy Directions*

Consider an arrangement of n equivalent easy directions distributed
in a plane P, such that the angle between any two neighbouring directions
is constant and equal to $2\pi/n$.   Also consider a magnetic field **H**, suffi-
ciently large for saturation, to be applied at an angle α to the normal
to P, such that the component of **H** in P makes an angle θ with a particular
easy direction A in P.   If $-\pi/n < \theta < \pi/n$ then after reducing **H** to zero
the magnetization vector $\mathbf{I}_0$ is parallel to A, the nearest easy direction.
Hence the reduced remanent magnetization is given by

$$j_r(\theta, \alpha) = \cos \theta \, \sin \alpha, \quad |\theta| < \pi/n. \qquad . \quad . \quad . \quad (2.1)$$

For an assembly of particles with a random distribution of easy directions
the reduced remanent magnetization is given by

$$j_r = \int_0^{\pi/2} \int_0^{\pi/n} \cos \theta \, \sin^2 \alpha \, d\theta \, d\alpha \bigg/ \int_0^{\pi/2} \int_0^{\pi/n} \sin \alpha \, d\theta \, d\alpha = \tfrac{1}{4} n \sin (\pi/n). \qquad (2.2)$$

Relation (2.1) could also be used in calculations for particles whose
easy directions have been partially aligned by a magnetic treatment ;

complete alignment ($\theta=0$, $\alpha=\frac{1}{2}\pi$) gives $j_r=1$. The values of j_r given by (2.2) are plotted in fig. 1; although they are only defined at integral values of n intermediate values have a physical significance if the condition of equivalence of the easy directions is relaxed, as in § 3. The applications of the results are given below.

Fig. 1

Reduced remanent magnetization. $j_r=I_r/I_0$: n, number of equivalent easy directions. Curve: regular planar distribution; circles: regular spatial distribution; horizontal lines: asymptotes.

2.2. *Spatial Distribution of Easy Directions*

If the n equivalent easy directions are regularly distributed in space the calculation of j_r proceeds by the method of Gans (1932) who gives values for $n=6$ and 8. In order to determine the trend of the j_r, n relation a calculation has been carried out for $n=4$, giving

$$j_r=(\sqrt{6}/\pi) \tan^{-1}(\sqrt{2})=0.7447. \qquad . \quad . \quad . \quad (2.3)$$

For $n=2$, $j_r=\frac{1}{2}$ as is also given by (2.2). The results for spatial distribution are shown by circles in fig. 1.

2.3. *Applications*

$n=1$. The remanent magnetization for a random assembly is zero. Unidirectional anisotropy has recently been discovered by Meiklejohn and Bean (1957) in cobalt particles with a cobalt oxide shell. A complete treatment of this case is possible and is given in the Appendix.

$n=2$. Uniaxial anisotropy.

$n=3$. In a calculation of the complete remanence curve for powders of haematite, $\alpha\text{Fe}_2\text{O}_3$ (Wohlfarth 1955 b), it was assumed that in the

easy plane of magnetization which occurs above the Morin transition temperature and lies perpendicular to the trigonal axis, anisotropy results from the effect of three coplanar easy directions of magnetization.   Several ferromagnetic resonance measurements on haematite (Anderson *et al.* 1954, Kojima 1955, Kumagai *et al.* 1955, Shimizu 1956) give differing indications about the nature and distribution of these easy directions, the values of n derived by the various authors ranging from $n=0$ to $n=6$.   Careful measurements of the remanent and saturation magnetizations should settle this question by reference to fig. 1.

$n=4$.   A planar distribution of four easy directions may occur in particles in which two types of uniaxial anisotropy act simultaneously, as discussed in § 3.

$n=6, 8$.   Gans' values for a spatial distribution apply respectively to spherical iron and nickel powders.

As found by Jonker *et al.* (1956, 1957) the powder particles of certain sintered mixed barium ferrites behave very similarly to haematite in having a difficult (hexagonal) axis of magnetization, with an easy plane lying perpendicular to it.   It is stated without detail that some anisotropy exists in this plane, and measurements of the remanent magnetization could again be used to give some information about the nature of this anisotropy.

The results shown in fig. 1 indicate that large values of j_r are obtainable even for randomly oriented particles if n, the number of easy directions, is large, and that a spatial distribution is more advantageous than a planar one.   Unfortunately it is not usually possible to develop magnetic materials with a pre-determined anisotropy, but the present results indicate what should be aimed at.   For spherical powders of iron n is indeed large, but here the magnetocrystalline anisotropy coefficient is relatively small, leading to a low coercive force (Néel 1947).   For practical applications iron powders are now produced which have a highly acicular shape ($n=2$), and j_r is increased by magnetic treatment (Paine *et al.* 1955 a, b, Mendelsohn *et al.* 1955).

§ 3. Mixed Uniaxial Anisotropies

In many instances the magnetic properties of particles are governed primarily by only one or other of the various types of uniaxial anisotropy which may arise.   Thus for iron particles having only slight deviations from spherical shape it is shape anisotropy which has the largest effect, while for cobalt or MnBi particles magnetocrystalline anisotropy predominates.   Recent measurements (see, for example, the work on dilute Ni–Au alloys by Berkowitz and Flanders 1957) have, however, shown that sometimes two anisotropies of not too unequal strengths may be active. Examples where these anisotropies are uniaxial are shape+magnetocrystalline anisotropy, shape+strain anisotropy, shape+shape anisotropy. The first case may arise, for example, for elongated hexagonal cobalt particles, and the last for non-ellipsoidal particles for which two long axes are clearly marked out.   An extensive programme of research on mixed

uniaxial anisotropies has been started, and the results for remanent magnetization, by far the easiest to calculate, are reported here.

3.1. *Anisotropy Energy Expressions*

Figure 2 shows schematically an arrangement of two unequal easy axes, characterized by anisotropy constants K_1 and K_2 and inclined at an angle β. In the remanent state the magnetization vector I_0 lies in the K_1, K_2 plane inclined at an angle ψ, say, with the K_1 axis. The anisotropy constants are to be regarded as general, and may be taken to

Fig. 2
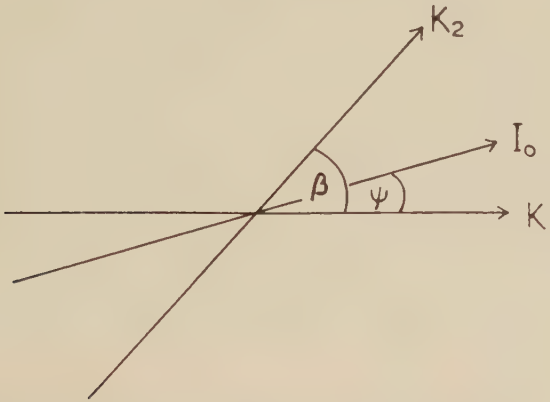
describe magnetocrystalline or shape or strain effects etc. For a single easy axis this generalization can always be made, and the same form of the anisotropy energy expression applies in all cases (Stoner and Wohlfarth 1948). In the present more complicated example this is no longer strictly true, and two different expressions for the anisotropy energy have been studied. These are, firstly,

$$E_A = K_1 \sin^2 \psi + K_2 \sin^2 (\beta - \psi), \qquad \text{. . .} \quad (3.1)$$

and, secondly,

$$E_A = K_1 \sin^2 \psi + K_2 \sin^2 (\beta - \psi) + \sqrt{(K_1 K_2)} \sin^2 \beta \sin^2 \psi \sin^2 (\beta - \psi). \quad (3.2)$$

The two expressions (3.1) and (3.2) have been derived by considering their forms in certain limiting cases, thus :

Relation (3.1). (*a*) If either $K_1 = 0$ or $K_2 = 0$, E_A reduces to a simple second order term appropriate to uniaxial anisotropy. (*b*) If $\beta = 0$ or π,

$$E_A = (K_1 + K_2) \sin^2 \psi, \qquad \text{. . . . . .} \quad (3.3)$$

again as appropriate to uniaxial anisotropy. (*c*) If $\beta = \pm \frac{1}{2}\pi$,

$$E_A = \text{const.} + (K_1 - K_2) \sin^2 \psi, \qquad \text{. . . .} \quad (3.4)$$

and, in particular, if $K_1 = K_2$, E_A is constant leading to anhysteretic behaviour. This state of affairs arises for a particle built up of two identical crossed ellipsoids. Although each of these gives rise to uniaxial

shape anisotropy the resulting anisotropy vanishes. This surprising result may be proved rigorously by using a theorem, derived independently by Brown and Morrish (1957) and by Rowlands (1956) and Rhodes and Rowlands (1957), and assuming that the compound particle is uniformly magnetized. The theorem states that for a particle of any shape the sum of the demagnetizing energies for uniform magnetization along any three perpendicular directions is constant and equal to $2\pi I_0^2$ per unit volume. The constancy of E_A is proved (we are indebted to Dr. P. Rhodes for informing us of this theorem and applying it to the present problem) by taking two of these directions to lie in the K_1, K_2 plane, the sum of the demagnetizing energies in the plane then being independent of ψ. But since here the two demagnetizing energies are by symmetry equal, each must be a constant. Although the internal forces (exchange etc.) for such a compound particle tend to keep the magnetization uniform, especially if the particle is not too large (Brown and Morrish 1957), the non-ellipsoidal shape necessarily gives rise to some (incalculable) degree of non-uniformity, especially if the particle is so small that thermal fluctuations become important (Néel 1949). Hence there will be some dependence of E_A on ψ even in this limiting case, and relation (3.1) is no longer strictly correct anywhere.

Relation (3.2). (*a*) If either $K_1=0$ or $K_2=0$, E_A reduces to a simple second order term as before. (*b*) If $\beta=0$ or π, (3.2) reduces to (3.3). (*c*) If $\beta=\pm\frac{1}{2}\pi$, then

$$E_A=\text{const.}+(K_1-K_2)\sin^2\psi+\sqrt{(K_1K_2)}\sin^2\psi\cos^2\psi, \qquad (3.5)$$

giving a second order as well as a fourth order term appropriate to the square symmetry. In particular, if $K_1=K_2=K$,

$$E_A=\text{const.}+K\sin^2\psi\cos^2\psi, \qquad \ldots \quad (3.6)$$

which is the correct expression if K denotes a magnetocrystalline anisotropy coefficient (in the usual expression involving direction cosines α_i, $\alpha_1=\sin\psi$, $\alpha_2=\cos\psi$, $\alpha_3=0$).

The relation (3.1) may thus be taken to apply primarily if the mixed anisotropy is of the type shape+shape, and (3.2) if it is of the type magnetocrystalline+magnetocrystalline. By obtaining and contrasting results based on both relations it is possible to describe with fair certainty any actual physical situation, such as shape+magnetocrystalline, which will usually be intermediate in its consequences between the two situations described by (3.1) and (3.2).

It may be noted that a more general relation for the anisotropy energy may be written

$$E_A=K_1\sin^2\psi+K_2\sin^2(\beta-\psi)+\eta\sqrt{(K_1K_2)}\sin^2\beta\sin^2\psi\sin^2(\beta-\psi), \quad (3.7)$$

where η is a constant. Relation (3.7) reduces to (3.1) and (3.2) if $\eta=0, 1$ respectively. With an additional parameter it would, however, be difficult to draw any very general conclusions from the results of the necessarily lengthy calculations.

3.2. *Remanent Magnetization for Random Particles*

(a) *Relation* (3.1)

In the remanent state the magnetization vector $\mathbf{I_0}$ lies in the K_1, K_2 plane along an energy minimum direction nearest to the direction of the field, previously applied at an angle α to the normal to this plane. Now relation (3.1) may be written, apart from a constant,

$$E_A = -\tfrac{1}{2}\{K_1{}^2 + K_2{}^2 + 2K_1 K_2 \cos 2\beta\}^{1/2} \cos 2(\psi - \xi), \qquad (3.8)$$

where $\xi = \tfrac{1}{2} \tan^{-1}\{K_2 \sin 2\beta/(K_1 + K_2 \cos 2\beta)\}$. Hence in the range $0 \leqslant \psi \leqslant \pi$ an energy minimum occurs at $\psi_m = \xi$ and a maximum at $\psi_M = \tfrac{1}{2}\pi + \xi$, with corresponding turning points in the range $\pi \leqslant \psi \leqslant 2\pi$. For a random assembly the reduced remanent magnetization is given by

$$j_r = \int_0^{\pi/2} \int_0^{(\psi_M - \psi_m)} \cos \theta \sin^2 \alpha \, d\theta \, d\alpha \bigg/ \int_0^{\pi/2} \int_0^{\pi/2} \sin \alpha \, d\theta \, d\alpha$$

$$= \tfrac{1}{2} \sin (\psi_M - \psi_m). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3.9)$$

Here, therefore, $j_r = \tfrac{1}{2}$ for all values of K_1, K_2 and β, and j_r retains this value even for an assembly of particles for which β varies in any fashion. The assembly thus behaves in this respect (in no other) as if all the particles were uniaxial.

Fig. 3
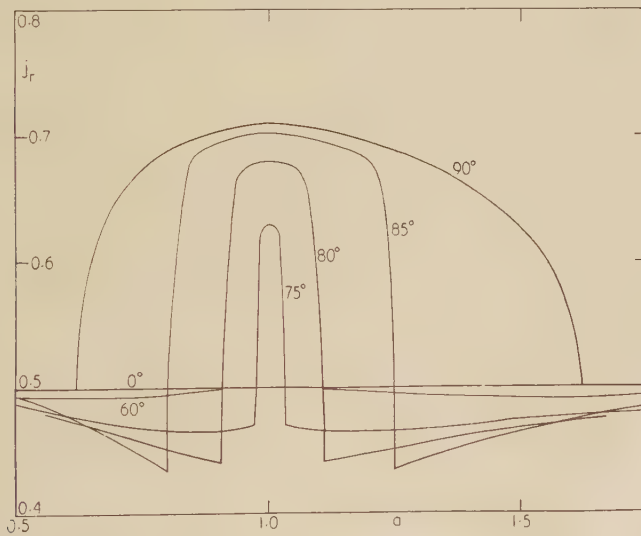
Reduced remanent magnetization, random assembly. $j_r = I_r/I_0$; $a = \sqrt{(K_1/K_2)}$; K_1, K_2, anisotropy coefficients; β, angle between easy axes, shown. Relation (3.2).

(b) *Relation* (3.2)

Here the distribution of energy maxima and minima in the K_1, K_2 plane is more complicated and has to be derived numerically. It is found that if β, the angle between the two easy axes, is less than a critical

value $\beta_c = 71 \cdot 87°$, then in the range $0 \leqslant \psi \leqslant \pi$ there is only one energy minimum, at ψ_m, and one maximum, at ψ_M. Here $(\psi_M - \psi_m)$ is not in general equal to $\frac{1}{2}\pi$ and j_r, given by (3.9), is less than $\frac{1}{2}$. If $\beta > \beta_c$, then for a range of values of the convenient parameter $a = \sqrt{(K_1/K_2)}$ there are two energy minima, at ψ_{m1} and ψ_{m2}, and two maxima, at ψ_{M1} and ψ_{M2}, such that $\psi_{m1} < \psi_{M1} < \psi_{m2} < \psi_{M2}$. The reduced remanent magnetization is then given by

$$j_r = \tfrac{1}{4}\{\sin(\psi_{M1} - \psi_{m1}) + \sin(\psi_{M2} - \psi_{m1}) + \sin(\psi_{m2} - \psi_{M1}) + \sin(\psi_{M2} - \psi_{m2})\}$$
$$. \quad . \quad . \quad (3.10)$$

and may increase to values above $\frac{1}{2}$. For $\beta = \frac{1}{2}\pi$ and $a = \sqrt{(K_1/K_2)} = 1$, $j_r = 2^{-1/2}$, as given in fig. 1 ($n = 4$, planar distribution). Figure 3 shows the dependence of j_r on $a = \sqrt{(K_1/K_2)}$ for a range of values of β. It is seen that, except for values of β close to $\frac{1}{2}\pi$ and of a close to 1, j_r does not deviate appreciably from $\frac{1}{2}$. Hence if the assembly contains particles for which the angle β varies randomly the mean value of j_r must be close to $\frac{1}{2}$ for all values of the parameter a. The largest deviation occurs at $a = 1$ where $j_r = 0 \cdot 550$.

Both the energy expressions (3.1) and 3.2) thus lead to almost the same conclusion, that the remanent magnetization of a random particle assembly is changed relatively little even if the anisotropy is more complicated than uniaxial and is of the mixed type considered here.

3.3. *Remanent Magnetization for Partially Aligned Particles*

It may happen that in an assembly of particles with mixed uniaxial anisotropies one set of easy axes, K_1 say, has been aligned by a magnetic treatment, while the other set, K_2, has not been aligned and is therefore randomly oriented. Hence the reduced remanent magnetization decreases below the ideal value $j_r = 1$ for complete alignment of the K_1 axes. In this section this decrease is determined as a function of the parameter $a = \sqrt{(K_1/K_2)}$.

(*a*) *Relation* (3.1)

From (3.8), assuming complete alignment of the K_1 axes and randomness of the K_2 axes,

$$j_r(a) = \int_0^{\pi/2} \cos \xi(\beta, a) \sin \beta \, d\beta. \quad . \quad . \quad . \quad (3.11)$$

By reducing (3.11) to

$$j_r(a) = \frac{1}{4a^2} \int_{|1 - a^2|}^{1 + a^2} \left\{ \frac{y(y + a^2 + 1)}{(y - a^2 + 1)} \right\}^{1/2} dy, \quad . \quad . \quad (3.12)$$

which is expressible in terms of elliptic integrals, or by expressing $j_r(a)$ in the form of series which converge rapidly except when $a \sim 1$, the values of j_r may be calculated and are shown in fig. 4. It is seen that for $a > 1$, i.e. $K_2 < K_1$, j_r does not deviate appreciably from 1, i.e. the influence of the randomly distributed set of easy axes K_2 does not greatly

' spoil ' the effect on the remanent magnetization of the alignment of the K_1 axes. The deviation become large, however, if $a < 1$, i.e. $K_2 > K_1$, and for $a = 0$, j_r is obviously equal to $\frac{1}{2}$.

(b) *Relation* (3.2)

The calculations are more difficult and the value of j_r has only been obtained at $a = 1$, where it is equal to $0 \cdot 927$ (shown by a circle in fig. 4) as compared with $j_r = 0 \cdot 862$ from relations (3·1) and (3·12). The general conclusion, that the remanent magnetization is not decreased very greatly if $a > 1$, thus holds even more strongly if the calculations are based on the relation (3.2).

Fig. 4
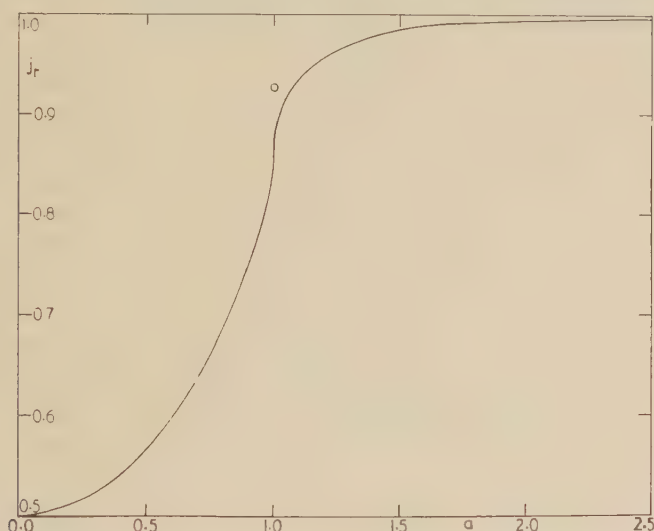
Reduced remanent magnetization, partially aligned assembly. $j_r = I_r/I_0$; $a = \sqrt{(K_1/K_2)}$; K_1, K_2, anisotropy coefficients ; K_1 axes aligned, K_2 random. Curve, relation (3.1) ; circle, relation (3.2).

§ 4. Conclusions

The work described shows that unless the number of easy directions of magnetization in a particle (more generally the number of energy minima) can be increased well above 2, the remanent magnetization of a random assembly is difficult to increase appreciably above $\frac{1}{2}I_0$. For an assembly of aligned uniaxial particles, however, even the complicating effect of an additional random uniaxial anisotropy, if not too strong, does little to decrease the remanent magnetization below I_0. The work described is now being extended to other properties such as susceptibility and coercive force. It should also be possible to discuss more complicated mixed anisotropies, such as uniaxial shape + cubic magnetocrystalline.

APPENDIX

Unidirectional Anisotropy

Meiklejohn and Bean (1957) have shown that for spherical single domain cobalt particles covered with an antiferromagnetic CoO shell the exchange interaction between the two phases is essentially unidirectional. They gave a brief discussion of the magnetization curve when the easy direction is parallel to that of the applied field. Here a simple and more general discussion of the magnetization curve is given.

Fig. 5
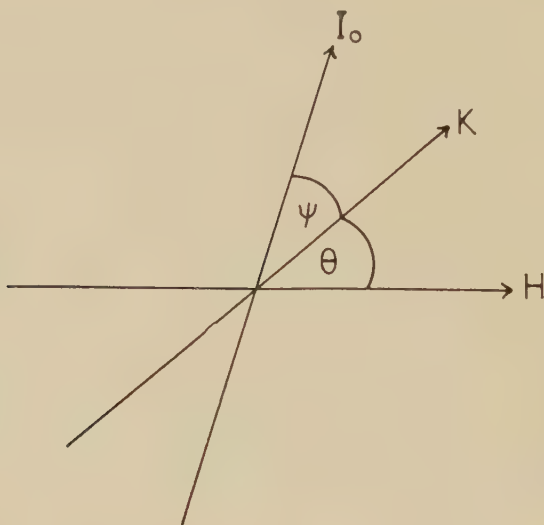
Figure 5 shows schematically the easy direction, characterized by an anisotropy constant K and making an angle θ with the positive direction of the applied field **H**. The angle between the magnetization vector I_0 and the easy direction is denoted by ψ. The total energy may be written, apart from a constant,

$$E = E_H + E_A = -HI_0 \cos(\theta + \psi) + K \sin^2(\tfrac{1}{2}\psi), \quad . \quad . \quad (A.1)$$

where $K = 2K_u$ in Meiklejohn and Bean's notation. For a minimum

$$h \sin(\theta + \psi) + \sin\psi = 0, \text{ i.e. } \psi = \tan^{-1}(1 + h \cos\theta)/h \sin\theta) - \tfrac{1}{2}\pi, \quad (A.2)$$

where $h = 2HI_0/K$. Hence the reduced magnetization is given by

$$j(h, \theta) = I/I_0 = \cos(\theta + \psi) = (h + \cos\theta)(h^2 + 2h \cos\theta + 1)^{-1/2}. \quad (A.3)$$

The process of magnetization is reversible except when $\theta = 0, \pi$, when discontinuous rotations of the magnetization vector take place at $h = \mp 1$ respectively; even here the j, h relation is single-valued. The asymmetrical curve for $\theta = 0$ is shown as curve 1 in fig. 6; this corresponds to an assembly of particles aligned by a magnetic field, as described by Meiklejohn and Bean.

Fig. 6
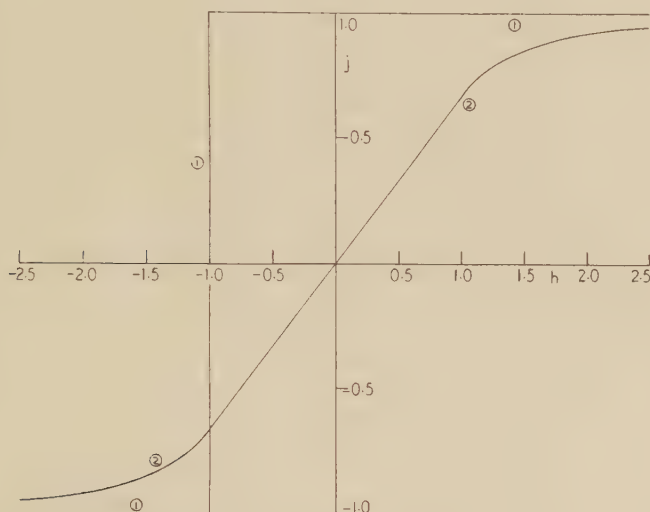
Magnetization curves for unidirectional anisotropy. $j = I/I_0$; $h = 2HI_0/K$.
Curve 1, aligned; 2, random.

For a random particle assembly the reduced magnetization is given by

$$j(h) = \tfrac{1}{2} \int_0^\pi j(h, \theta) \sin\theta \, d\theta = \left.\begin{aligned} &= -(1 - \tfrac{1}{3}h^{-2}), && h \leqslant -1 \\ &= \tfrac{2}{3}h, && -1 \leqslant h \leqslant 1 \\ &= 1 - \tfrac{1}{3}h^{-2}, && h \geqslant 1 \end{aligned}\right\} \quad . \quad (A.4)$$

The curve given by (A.4) is shown as curve 2 in fig. 6; the curve is symmetrical with zero remanent magnetization, as shown in fig. 1 ($n=1$).

For an assembly of cobalt + CoO particles exhibiting this unidirectional anisotropy as well as the magnetocrystalline anisotropy of cobalt, the resultant effect of the mixed anisotropies could be discussed by the methods of this paper. The resultant hysteresis loop of the particles is observed to be displaced along the H axis if the particles are aligned magnetically (see Meiklejohn and Bean 1957, fig. 1). This is easily explained (as shown by these authors) by combining the curve 1 of fig. 6 with the symmetrical loop for cobalt. It is unlikely, however, as discussed in this paper, that the two anisotropies always act in such a simple additive fashion. If the particles are not aligned, the resultant loop is observed

to be symmetrical with extended straight sides (q.v., fig. 1).   Such a loop is obtained by combining the curve for random unidirectional anisotropy (fig. 6, curve 2), which has an extended straight portion, with the symmetrical loop for cobalt.

REFERENCES

ANDERSON, P. W., MERRITT, F. R., REMEIKA, J. P., and YAGER, W. A., 1954, *Phys. Rev.*, **93**, 717.

BERKOWITZ, A. E., and FLANDERS, P. J., 1957, *Proc. Boston Conf.*, p. 275.

BROWN, W. F., and MORRISH, A. H., 1957, *Phys. Rev.*, **105**, 1198.

GANS, R., 1932, *Ann. Phys., Lpz.*, **15**, 28.

JONKER, G. H., WIJN, H. P. J., and BRAUN, P. B., 1956, *Philips tech. Rev.*, **18**, 145 ;   1957, *Proc. Inst. Elect. Engrs.*, **104**, 249.

KOJIMA, Y., 1955, *Sci. Rep. Res. Insts. Tôhoku Univ.*, **7**, 591.

KUMAGAI, H., ABE, H., ÔNO, K., HAYASHI, I., SHIMADA, J., and IWANAGA, K., 1955, *Phys. Rev.*, **99**, 1116.

MEIKLEJOHN, W. H., and BEAN, C. P., 1957, *Phys. Rev.*, **105**, 904.

MENDELSOHN, L. I., LUBORSKY, F. E., and PAINE, T. O., 1955, *J. appl. Phys.*, **26**, 1274.

NÉEL, L., 1947, *C. R. Acad. Sci., Paris*, **224**, 1488 ;   1949, *Ann. Géophys.*, **5**, 99.

PAINE, T. O., MENDELSOHN, L. I., and LUBORSKY, F. E., 1955 a, *Phys. Rev.*, **100**, 1055 ;   1955 b, *Belg. Pat.*, 546, 756.

RHODES, P., and ROWLANDS, G., 1957, (to be published).

ROWLANDS, G., 1956, *Thesis*, University of Leeds.

SHIMIZU, M., 1956, *J. phys. Soc., Japan*, **11**, 1078.

STONER, E. C., and WOHLFARTH, E. P., 1948, *Phil. Trans.*, **240**, 599.

WOHLFARTH, E. P., 1955 a, *Research*, **8**, S 42 ;   1955 b, *Phil. Mag.*, **46**, 1155 ; 1956, *Sci. J. R. Coll. Sci.*, **26**, 19.

The Solution of Inert Gas Atoms in Metals †

By D. E. Rimmer and A. H. Cottrell

Atomic Energy Research Establishment, Harwell

Abstract

Estimates are made of the energies of interstitial and substitutional solutions of inert gas atoms in a simple metal. The loss of cohesion round such an atom is taken into account by regarding the atom as occupying a vacancy in the metal. Most of the energy of solution, particularly for the larger atoms, originates from the closed-shell interactions, which are estimated from interaction potentials based on data for the inert gases and on Huntington's potential for copper ions. Argon, krypton and xenon are expected to dissolve substitutionally in copper under all circumstances, even to the extent of forcing copper atoms off the lattice if no vacant sites are available. Neon should and helium may dissolve substitutionally when vacancies are available, but not otherwise. The strains round xenon and krypton atoms are so large that extra vacancies may be captured by these atoms to reduce the strain energy. The solubilities and diffusion of inert gas atoms in metals are briefly discussed in the light of these conclusions.

§ 1. Introduction

HELIUM is produced by nuclear reactions in some materials used in nuclear reactors ; for example, beryllium and boron. Krypton and xenon are produced in uranium by fission. The solutions of these inert gas atoms in solids are interesting. Extremely high supersaturations can be obtained when the irradiations are made at temperatures too low for atomic diffusion to take place; and at high temperatures the atoms often escape from solution into gas bubbles, causing some materials to swell and some to break up.

In this paper we shall discuss the solution of a single inert gas atom in a metallic crystal. There are two problems :

(i) Let the atom be placed interstitially in a region of perfect lattice (having arrived there, for example, as a fast ion). We have to decide whether the atom will remain in interstitial solution, or go into substitutional solution by pushing one or more neighbouring atoms off the lattice into interstitial positions. Such states of solution are in general not equilibrium states, but are important at low temperatures where

† Communicated by the Authors.

thermal vacancies are not available. The ability of the gas to escape
from solution at low temperatures depends, for example, on its ability
to diffuse interstitially, and we might expect a small atom such as helium
to behave quite differently in this respect from a large one such as xenon.

(ii) As (i), but let the metal be heated to provide thermal vacancies.
The atom can now reach the equilibrium state of solution by capturing
vacancies and re-arranging metal atoms round itself quite freely. The
energy of this state determines the solubility of the gas and the atomic
configuration of the state determines the atomic mobility.

§ 2. The Energy of Solution

We consider the two main contributions to the energy of solution to
be the strain energy in and round the gas atom, and the loss of cohesive
energy at the site of the atom. The van der Waals' attraction between
the atom and its neighbours is not considered separately but is included
in the atomic interaction potentials from which the strain energy is to be
estimated. Other contributions are expected to be small ; the zero-point
vibrational energy, for instance, is not expected to exceed about 0·1 ev
even in the extreme case of helium dissolved interstitially in beryllium.

Because the ionization and resonance potentials of the inert gases are
so high (10 to 25 ev) we expect the electrons of such an atom to remain
in atomic orbitals and not to contribute appreciably to cohesion. The
atom can thus be regarded, when in substitutional solution, as occupying
a vacancy in the metal, and we shall take the energy of a vacancy as
a measure of the loss of cohesive energy associated with the atom. This
approach is fairly obvious when the gas atom occupies a substitutional
site. But it has also to be used for interstitial solutions when the volume
is increased by the presence of the atom.

Following the method of Friedel (1954), Fumi (1955) has shown that
the energy of a vacancy in a monovalent metal is contributed mainly
by the conduction electrons. The removal of a positive ion to create
the vacancy is equivalent to the addition of a localized potential which
repels locally a unit charge of conduction electrons. The corresponding
increase, $\frac{2}{3}E_F$, in energy, where E_F is the width of the conduction band,
is partly offset by the change in Fermi energy, $-\frac{2}{5}E_F$, due to the expansion
of the metal by one atomic volume. Thus, in this approximate theory,
the energy to form the vacancy is $\frac{4}{15}E_F$; in a refinement of this calculation
Fumi (1955) obtains the value $\frac{1}{6}E_F$, which is about 1 ev for copper, silver
and gold.

The same theory can be applied to the inert gas atom in solution.
The positive ions are either removed (substitutional solution) or repelled
(interstitial solution) to make room for the atom, and the ensuing charge
shift of conduction electrons requires about $\frac{4}{15}E_F$ of energy for each
atomic volume taken up by the atom. The strained lattice round the
gas atom is dilated and (except in calculations where the effect of this is

included implicitly, as for example in elasticity theory) the change in energy of the conduction electrons due to this dilatation must also be included. Here again, to the same level of approximation, the energy of the conduction electrons increases by $\frac{4}{15}E_F$ for each atomic volume of expansion (Cottrell *et al.* 1953).

§ 3. The Lattice Sum Method

We shall estimate the strain energy by a lattice sum method similar to that used by Huntington (1953) in the study of interstitial lattice defects. First, we evaluate the mutual potential energy of closed-shell interactions for gas–metal and metal–metal pairs of atoms. The strain energy of a configuration can then be found and the positions of certain atoms can be adjusted until the total strain and cohesive energy becomes a minimum.

The only metal for which a mutual potential energy is known with any confidence is copper, and all our calculations therefore make use of the potential given by Huntington (1953) for this metal,

$$U(r) = 0{\cdot}053 \exp\left[13{\cdot}9(r_0-r)/r_0\right] \quad . \quad . \quad . \quad . \quad (1)$$

ev per ion pair, where r is the actual and r_0 the equilibrium ($2{\cdot}56$ Å) separation of the ions. This is the ' softer ' of the two potentials used by Huntington, and may cause the strain energy to be underestimated, but it appears to be the more accurate of the two. It excludes the effect of the electron gas and so we have to take the total dilatation produced by the gas atom into account when estimating the energy of the conduction electrons.

To estimate the interaction potential of a gas-copper pair of atoms we have to combine the above potential with one for the gas atom. Let r be the distance between the centres of the gas and copper atoms, r_g be the radius of the gas atom and $r_c = (r-r_g)$ that of the copper atom. We assume that the interaction potential between these atoms can be written in the form

$$\epsilon_{gc}(r) = \tfrac{1}{2}\left[\epsilon(2r_g) + U(2r_c)\right]. \quad . \quad . \quad . \quad . \quad (2)$$

where $U(2r_c)$ is given by eqn. (1) and $\epsilon(2r_g)$ is the mutual potential of two gas atoms, and minimize $\epsilon_{gc}(r)$ with respect to r_g and r_c for a fixed r. This minimum value, as a function of r, is taken to be the interaction potential.

The interatomic potentials for the inert gases are not known accurately for the separations involved here, except for helium and neon where they have been derived theoretically. In the other cases the interaction must be obtained from compressibility data and transport properties, measured under conditions where the atoms do not approach as closely as in the present problem. Repulsive potentials have in fact been determined at very small separations from the scattering of fast ions, cf. Amdur *et al.*

(1950), Amdur and Mason (1954) and Berry (1949), but in these two-body collisions there is a possibility of lowering the energy by a redistribution of charge round the ions which cannot exist for an atom compressed in all directions, as in a solid.

To meet these difficulties, two different potentials are used for argon, krypton and xenon, since the differences between these should provide an indication of the errors to be expected from inaccurate potentials. For argon, the first is the potential

$$\epsilon_A(r) = 10\ 550 \exp(-r/0.273) - 63.8r^{-6} \quad \cdots \quad (3)$$

ev (r in Å) derived by Buckingham (1938) from the equation of state. The second is

$$\epsilon_A(r) = 1270 \exp(-r/0.334) - 79.7r^{-6}, \quad \cdots \quad (4)$$

obtained from the compressibilities measured by Stewart (1956) up to 19 000 kgm/cm². This falls below Buckingham's curve at the separations involved here. For krypton and xenon, Mason and Rice (1954) have derived the potential

$$\epsilon(r) = \frac{\beta}{1 - (6/\alpha)} \left[\frac{6}{\alpha} \exp\left\{ \alpha\left(1 - \frac{r}{r_m}\right) \right\} - \left(\frac{r_m}{r}\right)^6 \right] \quad \cdots \quad (5)$$

from a range of available data, where

	α	r_m (Å)	β/k (°K)
Kr	12·3	4·056	158·3
Xe	13·0	4·450	231·2

and k is Boltzmann's constant. As alternatives the potentials

$$\epsilon_{Kr}(r) = 2.74 \times 10^5 r^{-12} - 128r^{-6}, \quad \cdots \quad (6)$$

$$\epsilon_{Xe}(r) = 1.57 \times 10^6 r^{-12} - 349r^{-6}, \quad \cdots \quad (7)$$

obtained by Beattie *et al.* (1951, 1952) from compressibility measurements have been used. These are larger than those of Mason and Rice in the range considered here.

We can try to verify these potentials from data on alkali halides, on the assumption that the repulsive interaction of two inert gas atoms is the same as that between alkali metal and halogen ions with identical electronic configurations. Thus argon is compared with KCl, krypton with RbBr, and xenon with CsI. The electrostatic attractions between these ions enable the repulsive potential to be measured at a smaller separation than can be achieved in inert gases. The repulsive energy can be evaluated at the equilibrium separation in the solid, provided the form of the repulsive potential is known. On the whole, a 12th power repulsion gives better agreement with the inert gas values than do the more frequently used 9th power and exponential forms, as the results in table 1 show.

It is seen that the energies estimated from the ionic data are generally higher than those obtained from the gas formulae, but the individual values vary from 70% above them to 10% below them.

The potential we use for helium is

$$\epsilon_{He}(r) = 481 \exp(-r/0.218) - 0.93r^{-6}, \quad \ldots \ldots (8)$$

derived theoretically by Slater and Kirkwood (1931), and that for neon is

$$\epsilon_{Ne}(r) = 1605 \exp(-r/0.235) - 5.62r^{-6}, \quad \ldots \ldots (9)$$

derived by Buckingham (1938) from experimental data.

Table 1. Comparison of Repulsive Energies obtained from Gases and Ionic Crystals

Gas	Alkali Halide, MN	Equilibrium spacing of MN ions (Å)	Repulsive interaction (ev) at equilibrium spacing				
			Gas formulae		Ionic crystals		
			(i)	(ii)	9th	12th	Exponential
A	KCl	3·15	0·101	0·102	0·151	0·113	0·140
Kr	RbBr	3·45	0·098	0·080	0·139	0·104	0·127
Xe	CsI	3·95	0·107	0·074	0·125	0·094	0·097

(i) refers to the values of Buckingham for A and of Beattie *et al.* for Kr and Xe.
(ii) refers to the values of Stewart for A and of Mason and Rice for Kr and Xe.

With interaction potentials available for all pairs of atoms in the lattice the energy of any configuration is uniquely determined. For the interstitial gas atom the method of minimizing the energy is based on the first part of Huntington's calculation of an interstitial defect in copper (Huntington 1953). The six neighbours of the atom are moved away from it by a distance $\frac{1}{2}xd$ where d is the lattice spacing, and x is a variable. This lowers the gas–copper energy but decreases the distance between certain copper atoms. It is therefore necessary, for each of the six atoms already moved, to move a further four atoms away by a distance $\frac{1}{2}yd$, where y is another variable. Adjustment of these thirty atoms reduces the energy to a value that must be near the true minimum. First, the values of y which minimize the energy for a fixed x are determined as a function of x. This energy is then minimized with respect to x.

The substitutional case is simpler. As there is less lattice distortion only one relaxation parameter, λ, which gives the distance moved outwards by the twelve neighbours of the gas atom, is used in minimizing the energy.

The values of y and λ so obtained can be used to estimate the dilatation and electronic energy. This is most easily done by making comparisons with some known figures, as follows :

(i) As explained in § 2, the increase in electronic energy, in ev, is approximately equal to the dilatation measured in atomic volumes.

(ii) The dilatation associated with a copper interstitial is about two atomic volumes (Lomer 1957).

(iii) The effective linear strain, from which the dilatation is obtained, due to n uniformly distributed centres of pressure per unit volume is

$$e(n) = (4\pi/3)nc\gamma \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

(Tucker and Sampson 1954) ; c is defined by $u = c/r^2$, where u is the displacement at a distance r from the centre of pressure, and γ is a function of the elastic moduli of the lattice and of the misfitting atom, but may be taken as constant for a given lattice.

Hence by comparing the value of y with that of Huntington for displacements at such distances from the interstitial atom that the strains are small, the relative dilatation for the interstitial solution may be found. Using (i) and (ii) we then obtain the corresponding increase in electronic energy, which is added to the interatomic sum to give the energy of solution.

Similarly, we can use λ to obtain values of dilatation and electronic energy for the substitutional solution. But in this case we must add further terms. If the gas atom enters substitutional solution in a perfect lattice, as in the first problem of § 1, a vacancy–interstitial pair has to be created, which requires an energy $E_V + E_I$, the sum of their energies of formation. For copper we assume $E_V \simeq 1$ ev and $E_I \simeq 4.5$ ev. If the atom enters substitutional solution by capturing a vacancy, as in the second problem of § 1, only the energy E_V of the vacancy has to be included.

§ 4. Results and Discussion

The results are given in table 2, in which E_i, E_{S1}, and E_{S2}, are the estimated total energies, in ev, for the interstitial and the two types of substitutional solution, respectively. We see that argon, krypton and xenon, are expected to dissolve substitutionally under all circumstances.

Table 2. Estimated Energies (ev) of Solution of Inert Gases in Copper

Type of Solution	Energy terms	He	Ne	A (a)	A (b)	Kr (c)	Kr (d)	Xe (c)	Xe (d)
Interstitial	Strain	1·7	3·4	10·6	7·7	14·9	10·8	24·5	17·5
	Electronic	0·8	1·2	3·0	2·3	4·6	3·0	6·5	4·9
	E_i	2·5	4·6	13·6	10·0	19·5	13·8	31·0	22·4
Substitutional (problem 2)	Strain	0	0·3	2·2	1·7	3·8	2·7	7·8	5·7
	Electronic	0	0·1	0·7	0·5	1·4	0·8	2·4	1·5
	E_V	1·0	1·0	1·0	1·0	1·0	1·0	1·0	1·0
	E_{S2}	1·0	1·4	3·9	3·2	6·2	4·5	11·2	8·2
Substitutional (problem 1)	E_{S1} ($=E_{S2}+E_I$)	5·5	5·9	8·4	7·7	10·7	9·0	15·7	12·7

(a) Buckingham ; (b) Stewart ; (c) Beattie *et al.* ; (d) Mason and Rice.

Neon should certainly dissolve substitutionally when thermal vacancies are present ; the difference between E_i and E_{S1} is not large for this atom but is such as to favour interstitial solution when vacancies are absent. Helium should certainly dissolve interstitially when vacancies are absent, but may dissolve substitutionally when they are present.

These values can be compared with a relation derived by Lomer (1957) from Grüneisen's constant, the expansion coefficient and the compressibility of a metal. This relation shows that, for copper dilated by a finite shear strain, the strain energy, in ev, should be approximately 2·8 times the dilatation, in atomic volumes. The corresponding ratio in table 2 (remembering that the electronic term there has the same numerical value as the dilatation) varies from 1·9 to 3·6, increasing with the lattice strain. There is a rough agreement, and the differences may be due to inaccuracies in estimating the dilatation or to a breakdown of Lomer's relation at large strains.

The contribution of the electronic term to the energy of substitutional solution is not large and, by itself, it would not restrict the solubility of the inert gas atoms to abnormally low values. It must be remembered that although these atoms do not bond to the metal, they also do not bond to one another in the gaseous state. Thus the large molecular dissociation energy, which restricts the solution of so many gases in metals, and which exceeds 4 ev in the case of hydrogen, is absent for the inert gases. The fission gases krypton and xenon are expected to have extremely low solubilities in metallic crystals, not because of their lack of cohesion, but because of their great size, and the same conclusion is expected to hold for other fission products with similar ionic radii. In such cases, of course, some solubility may be observed in the liquid state where the strain energy will be largely eliminated.

The xenon atom is so large that a considerable strain energy exists even when it occupies a substitutional site of the lattice, and it seems likely that some change would occur to reduce this energy. One possibility is that the atom would capture one or more extra vacancies. The extra energy E_V (problem 2) or $E_V + E_I$ (problem 1) that this requires would have to be supplied from the following processes :

(i) Lattice atoms next to the gas atom have strained bonds. When one of them is replaced by a vacancy this strain energy is recovered.

(ii) Certain neighbours of this vacancy reduce their strain energy by moving towards it.

(iii) The decrease in lattice dilatation, mainly due to process (ii), decreases the electronic energy.

Only the first of these can be evaluated easily. A lower limit to the total energy change can be estimated by calculating (i). finding the best approximation to (ii) that can be obtained by relaxing a few atoms in crystallographic directions, and making an estimate of (iii) from these movements. An upper limit, which attempts to take account of more

radical rearrangements round the gas atom, can be estimated by assuming that the volume of the vacancy is spread uniformly round the gas atom. Table 3 gives term (i), i.e. the maximum energy that is certain to be recovered, and the estimated lower and upper limits for the larger gas atoms.   These figures show that a second vacancy is unlikely to be formed by the expulsion of a lattice atom into an interstitial position (problem 1), but that in thermal equilibrium xenon and possibly also krypton may capture extra vacancies.

Table 3.   Energies Released by the Addition of a Second Vacancy to a Gas Atom (in ev)

	Argon		Krypton		Xenon	
	Buckingham	Stewart	Mason	Beattie	Mason	Beattie
Term (i)	0·15	0·1	0·1	0·2	0·35	0·5
Lower limit	0·3	0·15	0·3	0·6	0·7	1·2
Upper limit	1·4	1·2	1·6	1·9	2·8	2·9

The upper limit could be approached much more easily if a few gas atoms were to associate together since the greater complexity of the configuration round such a cluster would make it possible to distribute the extra volume of each additional vacancy to the best effect.   There may thus be a sharp fall in the energy of solution when the gas atoms begin to cluster together, which would assist the process of nucleation of gas bubbles.

The association of vacancies with gas atoms should also cause these atoms to diffuse faster in the lattice.   Very low activation energies for diffusion of fission gases in uranium have been reported by several observers (Zimen and Schmeling 1954, Reynolds 1956, Zimen and Dahl 1957, Curtis and Rich 1957) but it seems that diffusion along grain boundaries and cracks may be predominant in these experiments.   LeClaire and Rowe (1955) obtained an activation energy of 33 600 cal/mol for the diffusion of argon in silver, and Tobin (1957) likewise obtained 35 000 cal/mol for krypton in silver.   These values are about 10 000 cal/mol below that for self-diffusion in silver, but the relatively large interatomic spacing in this metal (2·88 Å, compared with 2·56 Å for copper) is expected to make argon and krypton in this metal behave rather like neon and argon, respectively, in copper, and their ability to capture vacancies is correspondingly reduced.

References

AMDUR, I., DAVENPORT, D. E., and KELLS, M. C., 1950, *J. chem. Phys.*, **18,** 525.

AMDUR, I., and MASON, E. A., 1954, *J. chem. Phys.*, **22,** 670.

BEATTIE, J. A., BARRIAULT, R. J., and BRIERLEY, J. S., 1951, *J. chem. Phys.*, **19,** 1222.

BEATTIE, J. A., BRIERLEY, J. S., and BARRIAULT, R. J., 1952, *J. chem. Phys.*, **20,** 1615.

BERRY, H. W., 1949, *Phys. Rev.*, **75,** 913.

BUCKINGHAM, R. A., 1938, *Proc. roy. Soc.* A, **168,** 264.

COTTRELL, A. H., HUNTER, S. C., and NABARRO, F. R. N., 1953, *Phil. Mag.*, **44,** 1064.

CURTIS, G. C., and RICH, J. B., 1957 (to be published).

FRIEDEL, J., 1954, *Advanc. Phys.*, **3,** 446.

FUMI, F. G., 1955, *Phil. Mag.*, **46,** 1007.

HUNTINGDON, H. B., 1953, *Phys. Rev.*, **91,** 1092.

LeCLAIRE, A. D., and ROWE, A. H., 1955, *Revue de Métallurgie*, **52,** 94.

LOMER, W. M., 1957, *Phil. Mag.*, **2,** 1053.

MASON, E. A., and RICE, W. E., 1954, *J. chem. Phys.*, **22,** 843.

REYNOLDS, M. B., 1956, *Nuclear Sci. Engrg*, **1,** 374.

SLATER, J. C., and KIRKWOOD, J. G., 1931, *Phys. Rev.*, **37,** 682.

STEWART, J. W., 1956, *Phys. Chem. Solids*, **1,** 146.

TOBIN, J. M., 1957, *Acta Met.*, **5,** 398.

TUCKER, C. W. Jr., and SAMPSON, J. B., 1954, *Acta Met.*, **2,** 433.

ZIMEN, K. E., and DAHL, L., 1957, *Z. Naturf.*, **12a,** 167.

ZIMEN, K. E., and SCHMELING, P., 1954, *Z. Elektrochem.*, **58,** 599.

Indium-rich Indium–Magnesium and Indium–Lithium Alloys†

By J. Graham and G. V. Raynor

Department of Physical Metallurgy, University of Birmingham

Synopsis

For comparison with the results obtained from alloys of indium with solute metals from Group IIB of the periodic table, the lattice spacings of indium–magnesium and indium–lithium alloys have been measured. Although the results show an unexpected scatter, they establish that neither magnesium nor lithium cause a marked axial ratio decrease. A possible reason for this behaviour is briefly discussed.

The solid solubility of magnesium in indium at room temperature lies between 35 and 40 atm% of magnesium. The primary solid solution enters into equilibrium with the previously observed face-centred cubic intermediate solid solution (denoted β) above 275°c, and with the ordered form of the same intermediate phase below this temperature. There is a peritectic reaction $\beta + \text{liquid} \rightleftharpoons$ (indium-rich primary solid solution) at 333·5°c.

§ 1. Introduction

In the course of previous work in the authors' laboratory, the variation of lattice spacing with composition was determined for solid solutions in indium of cadmium, mercury, tin and lead (Tyzack and Raynor 1954). The effects observed were satisfactorily accounted for in terms of the variation, with change in electron : atom ratio, of the number of electrons per atom overlapping from the first Brillouin zone for the indium structure across the (002) planes of energy discontinuity. For this interpretation, it must be assumed that the solute metals from Groups IIB and IVB of the Periodic Table respectively contribute less electrons per atom and more electrons per atom than indium itself. Indium has frequently been described as possibly incompletely ionized in the solid state, and this makes quantitative analysis of the influence of electronic factors on lattice spacings difficult, especially since a similar tendency to incomplete ionization is present in the solute metals already referred to. It is of interest, therefore, to compare with previous results the lattice spacing changes produced in indium by solute metals free from the complication of incomplete ionization. Magnesium (valency 2) and lithium (valency 1)

are suitable solute metals for this purpose, and the present paper describes the results of experiments in the systems indium–magnesium and indium–lithium.

§ 2. Experimental Details

Exploratory alloys in the indium–magnesium system were prepared, in quantities of 1 to 4 g, from the following materials :

(i) ' Domal ' magnesium bar, 99·95% pure, obtained from Messrs. Magnesium Elektron Ltd.

(ii) ' Matthey ' indium wire, spectrographically standardized, obtained from Messrs. Johnson Matthey & Co. Ltd. Material was taken from three samples one of which contained 0·02% of lead as the major impurity, while a second contained 0·005% of iron ; the third sample contained trace impurities only, and was 99·998% pure.

The materials were weighed into Pyrex tubes, which were evacuated, flushed with argon, re-evacuated and sealed. The alloys were melted in an oxy-gas flame, well shaken, and the tubes quenched in cold water. Although spectrographical analysis confirmed that no silicon contamination had occurred, the resulting quenched ingots were dark in appearance, and tended to adhere to the glass. Further experiments were therefore made using ' H.S. ' magnesium rod obtained from Messrs. Johnson Matthey & Co. Ltd., which was 99·97% pure, the major impurity being 0·01% of iron. The method of preparation outlined above was then satisfactory. The absence of silicon contamination was again confirmed, but for purposes of comparison, several 4 g ingots were prepared, under conditions of continuous evacuation, in small stainless steel crucibles ; rapid chilling was obtained by quenching the crucibles and contents into cold water. Lattice spacing measurements were carried out only on the alloys for which ' H.S.' magnesium had been used.

In the indium–lithium system, for which high purity lithium containing less than 0·02% of sodium was employed, the method involving the use of Pyrex tubes proved entirely satisfactory.

After quenching, all ingots were severely deformed, re-sealed in evacuated Pyrex tubes, and annealed for at least ten days at approximately 125°c to ensure homogeneity. It was found that the indium–rich alloys recrystallized at room temperature ; diffusion was in general so rapid that quenching could not be relied upon to preserve high-temperature conditions. The homogenized ingots were therefore furnace cooled from the homogenizing temperature and allowed to stand for several days at room temperature before further treatment. X-ray specimens were then prepared by extrusion of a 0·4 mm diameter rod in a specially designed extrusion press. Analysis, where considered necessary, was carried out using the unextruded remainder. Strain relief of the extruded rod took place spontaneously at room temperature except for the more solute-rich alloys. After diffraction patterns had been obtained from solute-rich material, therefore, a vacuum-anneal at 100°c was given, followed by a second diffraction experiment after furnace cooling.

Difficulty was caused by the reactivity of the alloys in air, which increased with increasing solute content. After several trials, the most satisfactory method of protection proved to be the coating of rods, prepared as above, with a thin coating of silicone fluid.

All the indium–lithium alloys prepared were chemically analysed. The indium–magnesium alloys prepared in steel crucibles were also analysed, but since analysis of certain of those prepared in Pyrex gave results agreeing very closely with the intended compositions, the latter were in general accepted for material prepared by this method.

Specimens were exposed to cobalt K_α radiation in previously calibrated 9 cm Debye–Scherrer x-ray cameras ; the wavelengths accepted were :

$$CoK_{\alpha_1}\tfrac{1}{2}\lambda = 0 \cdot 89265 \text{ kx}$$
$$CoK_{\alpha_2}\tfrac{1}{2}\lambda = 0 \cdot 89459 \text{ kx.}$$

The films obtained were measured either with a Cambridge Universal Measuring Machine or a Hilger Vernier instrument ; the results from both instruments agreed. For the calculation of lattice spacings and the elimination of systematic errors, use was made of methods described by Cohen (1935, 1936) and by Archard (1953), as well as of the graphical method of adjusting the axial ratio to give the best straight lines when lattice spacings derived from various diffractions are plotted against the Nelson–Riley function $\sin^2 2\theta/\theta + \sin^2 2\theta/\sin\theta$ (Nelson and Riley 1945). The three methods gave almost indentical results, and for the majority of the experiments the speed and accuracy of Archard's method was preferred. For good films, the probable error in the a spacing, and in the axial ratio, are estimated to be 1 part in 20 000 ; for the c spacing, the probable error is 1 part in 10 000. The films obtained in the present work, however, were not of high quality and the errors were correspondingly greater. The results are expressed in kx units at 20°C ; the small corrections from the actual temperature of exposure were based upon the coefficient of thermal expansion of indium determined in previous work by the authors (Graham *et al.* 1955).

In the course of the work, certain indium–magnesium alloys were metallographically examined. Alloys of composition less than 5 atm% of magnesium were successfully electropolished in either of the following solutions :

(i) 1 vol. of glycerine, 2 vols. of pure perchloric acid, and 7 vols. of methyl alcohol.

(ii) 75 cm³ of carbitol, 5 cm³ of concentrated nitric acid, and 2 cm³ of concentrated hydrochloric acid. With care, alloys richer in magnesium could be hand-polished under paraffin, though it was difficult to preserve the polished surface sufficiently long for examination and photography.

§ 3. Previous Investigations

The indium–magnesium system has been previously examined by Hancke (1938) and by Raynor (1948), but there is little definite information with regard to the indium-rich alloys. According to Hancke, at

the equiatomic composition the crystal structure is body-centred tetra-
gonal with $a=3\cdot24$ and $c=4\cdot38$ kx : he interpreted an alloy containing
67 atm% of indium as ordered $MgIn_3$ displaced from its stoichiometric
composition, and stated that magnesium was soluble in indium though
no indication of the extent of the solubility was given. The work of
Raynor showed that, at high temperatures, the composition MgIn fell
within a wide range of face-centred cubic solid solutions, but that on
cooling alloys of approximately equiatomic composition an ordered
structure was formed. This was described as a superlattice of the
CuAu type, with an axial ratio of approximately $0\cdot96$; this description is
equivalent to that of Hancke if the atoms at the corners and centre of
the body-centred tetragonal cell are different. The form of the equi-
librium diagram for the indium–magnesium system from 0 to 50 atm%
of indium strongly suggested that the solid solubility of magnesium in
indium was very wide, extending from 100 atm% of indium to approxi-
mately 25 atm%, the tetragonal structure of indium transforming at
some composition above 50 atm% of indium to a cubic structure, as
in the case of the indium thallium alloys (Guttman 1950). This sugges-
tion was supported by metallographic and x-ray studies of one indium-
rich alloy containing $7\cdot34$ atm% of magnesium. This consisted of a
single phase tetragonal structure, for which the diffraction pattern
was similar to that of pure indium. The melting point of this alloy
lay between 175° and 180 c, and was consistent with extrapolation
of the solidus curve determined for alloys containing less than 50 atm% of
indium to the melting point of pure indium. It thus appeared unlikely
that the course of the solidus curve was interrupted by any reaction,
except possibly a peritectic reaction over a very narrow range of
composition, associated with the tetragonal $=$ cubic transformation. At
the outset of the present work, therefore, a wide solid solution of
magnesium in indium was expected.

The indium–lithium system has been investigated previously only by
Grube and Wolfe (1935), who reported some solubility of lithium in
indium, but the extent of the solubility was not defined.

§ 4. Results of Present Investigation

Fourteen binary indium–magnesium alloys were examined : the
compositions extended up to $45\cdot38$ atm% of magnesium, and the results
are plotted in fig. 1. There is considerably more scatter than usual
in these results. This is referred to in greater detail below.

The lattice spacings of pure indium at 20 c, in excellent agreement with
previous work, were found to be :

$$a=4\cdot5891\pm0\cdot0002\,\text{kx} ; \quad c=4\cdot9370\pm0\cdot0004\,\text{kx} ; \quad c/a=1\cdot0758.$$

Additions of magnesium up to 35 atm% gave rise to diffraction patterns
in which the sequence of reflections was visually indistinguishable from
that for pure indium ; there was however a marked increase in the general

background intensity, and the quality of the films deteriorated rapidly as the magnesium content increased.   No extra reflections were visible, suggesting that an extensive solid solution of magnesium in indium was formed ; this was confirmed by the recognition of extra diffraction lines due to a second constituent in the alloy containing 40·18 atm% of magnesium, corresponding to a body-centred tetragonal ordered structure of the distorted caesium–chloride type, with $a=3\cdot2476$ kx and $c=4\cdot4037$ kx†.   These measurements are in good agreement with the

Fig. 1
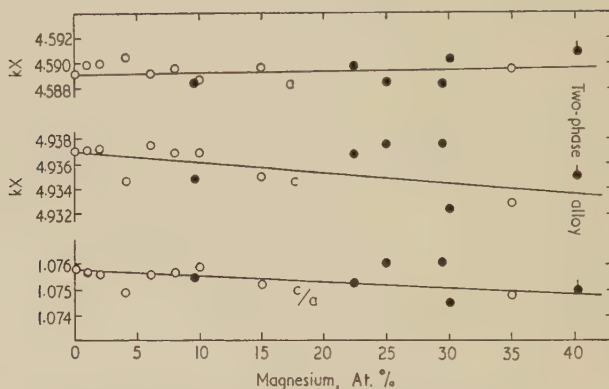
The lattice spacings and axial ratios for the solid solution of magnesium in indium.

previous work of Hancke (1938) and of Raynor (1948).   The solid solution of magnesium in indium thus extends at room temperature up to some composition between 35·04 and 40·18 atm% of magnesium, and then enters into equilibrium with the previously recognized ordered MgIn structure, which is face-centred cubic above the critical temperature for ordering (Raynor 1948).   The system therefore exhibits the typical tetragonal $\rightleftharpoons$ cubic transformation, which is a feature of indium-rich systems in which the solute metal has a group valency less than that of indium.

The significant feature of fig. 1, apart from the scatter of the experimental points, is the absence of any marked change of lattice spacing with composition.   Even at the solid solubility limit the lattice spacings are similar to those of indium.   The data plotted suggested that the *a* spacings increase very slightly with increase in magnesium content, while the *c* spacings decrease somewhat ; the resulting change in axial ratio is a slight decrease.

The scatter in the results makes it difficult to be certain of these general trends.   It is thought that the scatter is due to the difficulty of accurate

† This structure gave rise to the main pattern in the diffraction photograph of an alloy with 45·38 atm% of magnesium.

line measurement in the presence of relatively intense backgrounds. Numerous efforts were made to improve the quality of the films by heat-treatment of the x-ray specimens. but these were in general unsatisfactory, and the difficulty increased with increase in magnesium content.   In view of the slight dependence of lattice spacing on composition, and the rapidity of diffusion in these alloys, it is unlikely that the poor film quality was due to segregation or lack of equilibrium ; this is confirmed by the relative sharpness of the lines from the MgIn phase in the duplex alloys. It was, however, noted than on exposure to air of the ingots from which specimens were made, surface attack occurred, the rapidity of which again increased with increasing magnesium content.   The attack was sometimes so severe, in the grain boundaries of unprotected ingots, that the grains were forced apart ; very thin filaments, similar to the ' whiskers ' which have been recently studied by many workers (Compton *et al.* 1951), were left bridging the gap between grains.   These filaments were obtained in lengths up to 3 mm and qualitative tests showed them to have high elastic strength.   Previously, whiskers from indium have been grown only under pressure (Sines, private communication) ; the addition of as little as 1 atm% of magnesium, however, causes the spontaneous growth of filaments across the gap between grains forced apart by atmospheric attack.

Since the diffraction patterns obtained clearly showed the lines due to an indium-like structure. the effects observed could hardly have been due to selective oxidation of indium. leaving a surface layer rich in magnesium upon which the x-rays impinged.   Similarly the patterns could not be due to an indium-enriched layer left by selective magnesium oxidation, since in this case the systematic observation of second-phase diffraction patterns in the more magnesium-rich alloys of the series, together with patterns typical of only a very slightly altered indium structure, would have been unlikely.   It is probable. therefore, that the very slight change of lattice dimensions with composition is genuine, and that the poor quality of the films may be ascribed to the effect of the atmospheric attack on the protected specimen during exposure.   In view of this, it is doubtful how much weight should be ascribed to the lattice-spacing results for alloys containing more than 15 atm% of magnesium.

The results in general. however, suggest that a two-phase region exists at room temperature between the saturated solid solution of magnesium in indium and the ordered solid solution based on MgIn.   In order to investigate this more fully, alloys containing 37·59, 42·52 and 47·52 atm% of magnesium were prepared, using quantities of 5 g, in evacuated Pyrex tubes into which platinum/platinum–rhodium thermocouples had been sealed.   The liquid alloys were homogenized by vigorous shaking above the freezing point for 5 min ; differential thermal analysis was then carried out, using a similarly sealed zinc specimen as a standard.   Cooling and heating rates were approximately 1·3°c/min.   The results are plotted as open circles in the composite diagram of fig. 2.   Whereas the alloy with

37·59 atm% of magnesium gave only a liquidus arrest at 330°c, and a well-marked solidus arrest on heating at 285°c, the alloy with 42·52 atm% of magnesium gave a liquidus arrest at 353°c, followed by a constant temperature evolution of heat at 333·5°c ; no solidus arrest was observed, but an evolution of heat on cooling, and an absorption on heating, typical of an order–disorder reaction, was found, the critical temperature being 275°c. At 47·52 atm% of magnesium, the liquidus and solidus arrests occurred at 381° and 246°c respectively, while the critical temperature for ordering was 304°c. As shown in fig. 2, the solidus and order–disorder results are consistent with those previously obtained by Raynor

Fig. 2
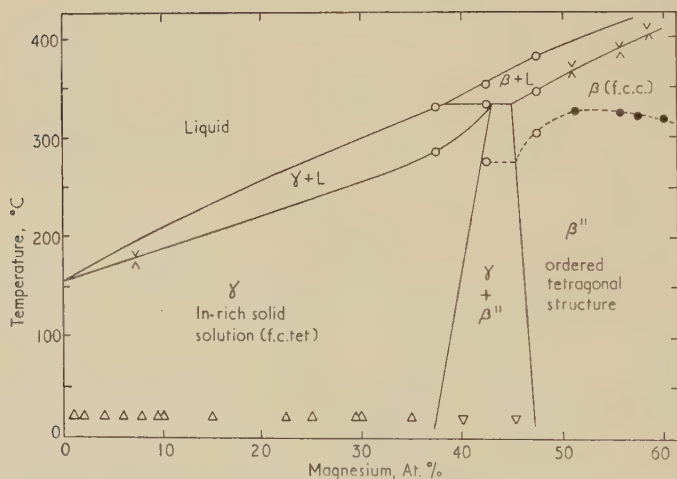
Probable form of the indium-rich portion of the indium–magnesium equilibrium diagram.

(1948) ; the present results indicate that the saturated indium–magnesium primary solid solution is formed peritectically from the disordered, face-centred cubic β solid solution, and that a two-phase region, which has been confirmed metallographically, extends over a narrow range of compositions between these solid solutions. At lower temperatures, the indium-rich solid solution is in equilibrium with the ordered tetragonal form of the β phase.

In contrast with the indium–magnesium system, the lattice spacings of indium are relatively strongly dependent on composition, as shown in fig. 3. The *a* spacing is increased, but the decrease in *c* as the percentage of lithium rises is sufficient to give rise to a decreasing axial ratio. There is again some scatter in the results, which is probably due to the reactivity of the alloys and the intense background on the films. All films showed

only the lines due to a face-centred tetragonal structure, and no lines due to a second phase were detected in the composition range examined.

Fig. 3
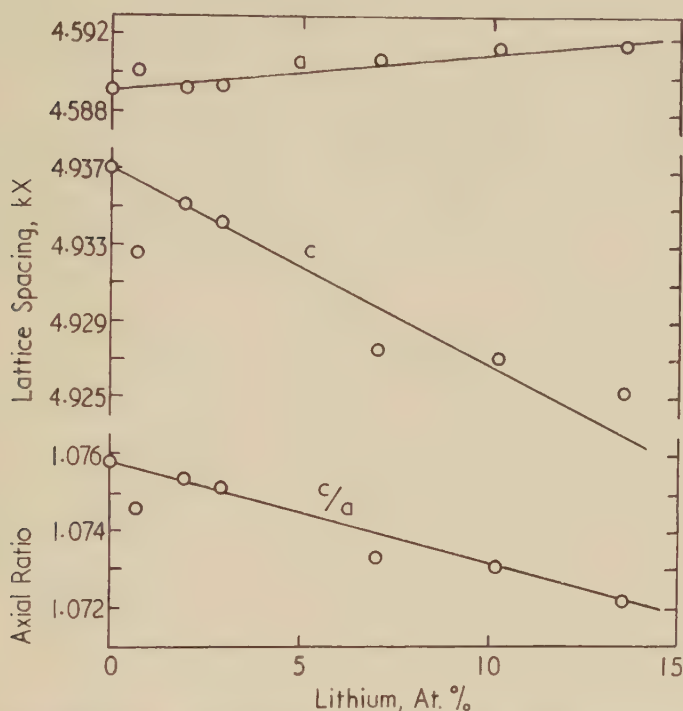
The lattice spacings and axial ratios for the solid solution of lithium in indium.

§ 5. Discussion

In previous work (Tyzack and Raynor 1954) it was pointed out that the effects of cadmium, mercury, tin and lead on the lattice spacings of indium could be interpreted satisfactorily in terms of the variation with composition of the overlap of electrons across the bounding planes of the first Brillouin Zone for the indium structure. The different magnitudes of the overlap across the four faces of types (200) and (020), and that across the two faces (002) was shown to be of particular importance. In general, according to this interpretation, a decrease in the magnitude of these overlaps due to the solution of a metal of effective valency less than that of indium, e.g. mercury or cadmium causes a relative contraction of the c axis of the crystal, with a consequent decrease in axial ratio.

The relative constancy of the lattice spacings of indium-rich indium–magnesium solid solutions is thus unexpected. Similarly, although the axial ratio of indium is decreased by the solution of lithium, the very small rate of change is difficult to understand. Thus, 10 atm% of lithium reduces the axial ratio to approximately 1·073, while the same reduction

is obtained by the addition of only 2 atm% of thallium, the effect of which is itself much less than that of the solutes cadmium and mercury referred to above.

The present work suggests that the Brillouin Zone effects are the predominant factor in the variation of lattice spacing with composition only for solute metals chosen from the B sub-groups of the Periodic Table and having considerable electronic and ionic similarity to indium.   Thus the ions of the elements cadmium, mercury, tin and lead (assuming complete ionization), all have an outermost group of ten 4d or 5d electrons ; the difference between cadmium and tin on the one hand and mercury and lead on the other lies mainly in the inclusion of the 4f group of 32 electrons within the ion.   Further the ionic radii are similar in magnitude to, or larger than, that of indium, as illustrated by the following values (Pauling 1927), expressed in Å :

In 1·04,   Cd 1·14,   Hg 1·25,   Sn 0·96,   Pb 1·06.

In the case of magnesium and lithium, however, the ionic radii are appreciably smaller than that of indium, being respectively 0·82 and 0·60 Å and the ions contain no d electrons.

An explanation of the difference between the effects of solutes like magnesium and lithium and those typified by cadmium and mercury may possibly be sought in terms of the general cohesion of the structure. If the cohesion in metallic indium involves bonding to which the 4d electrons contribute, then the substitution, for an atom of indium, of an atom of a solute metal which is also characterized by an ion with an outermost 4d or 5d group of electrons will leave this type of bonding relatively unaffected.   Changes in lattice dimensions will then depend mainly upon the relative atomic sizes of the components and the Fermi energy of the electrons which contribute to the first Brillouin Zone ; axial ratio changes will depend mainly upon the latter factor.   If however the solute element is characterized by an ion without d electrons (e.g. magnesium or lithium) then the cohesive forces in the solvent may well be decreased, leading to a general expansion of the structure.   Since the cohesive forces in indium appear to be smaller in the direction of the *c* axis than at right angles to this, the weakening in the structure would be expected to affect this direction predominantly, and an increase in the axial ratio would be expected if the weakening of the cohesion were the only factor involved.   The effect of the Fermi energy of the relatively free electrons, however, is to tend to decrease the axial ratio, just as in the indium–cadmium and indium–mercury solid solutions (Tyzack and Raynor 1954).   The very small changes in axial ratio produced by magnesium and lithium may thus be attributed to the opposition of these two factors.

The present work emphasizes that the effect on the lattice spacings of indium of solute metals taken from the short periods of the periodic table differs from that of B-subgroup elements ; further evidence of this is given by the fact that, whereas with mercury and cadmium the axial

ratio is decreased considerably before transformation to a cubic structure takes place, with magnesium as solute the transformation involves a cubic disordered phase, or at the appropriate temperature an ordered tetragonal phase, and an indium-rich solid solution of axial ratio almost indistinguishable from that of indium.

REFERENCES

ARCHARD, G. D., 1953, *Acta cryst.*, **6**, 813.
COHEN, M. U., 1935, *Rev. sci. Instrum.*, **6**, 68 ; 1936, *Ibid.*, **7**, 155.
COMPTON, K. G., MENDIZZA, A., and ARNOLD, S. M., 1951, *Corrosion*, **7**, 327.
GRAHAM, J., MOORE, A., and RAYNOR, G. V., 1955, *J. Inst. Metals*, **84**, 86.
GRUBE, G., and WOLFE, W., 1935, *Z. Elektrochem.*, **41**, 675.
GUTTMAN, L., 1950, *Trans. Amer. Inst. min. (metall.) engrs*, **188**, 1472.
HANCKE, W., 1938, *Naturwissenschaften*, **26**, 577.
NELSON, J. B., and RILEY, D. P., 1945, *Proc. phys. Soc. Lond.*, **57**, 160.
PAULING, L., 1927, *J. Amer. chem. Soc.*, **49**, 765.
RAYNOR, G. V., 1948, *Trans. Faraday Soc.*, **44**, 15.
TYZACK, C., and RAYNOR, G. V., 1954, *Trans. Faraday Soc.*, **50**, 675.

The Band Structure of the Transition Metals†

By N. F. Mott

Cavendish Laboratory, Cambridge

and K. W. H. Stevens‡

University of Nottingham

CONTENTS

Abstract

A discussion is given, based on the available results of experiment and of calculation, of the band structures of certain transition metals and of their relationship to their electrical and magnetic properties.

---

§ 1. Introduction

The purpose of this paper is to propose a model for the electronic structure of the transition metals based on recent experimental and theoretical work. In particular, we contrast the electronic structures of the body-centred metals chromium and iron with those of the close-packed metals cobalt and nickel.

The paper is partly based on an informal conference held in Cambridge on June 17 and 18, 1957, and reference is made to opinions expressed at this conference. It is realized that some of the conclusions reached are tentative, and some experiments are proposed which should shed light on their validity.

Some of the main conclusions reached are as follows. There is a very considerable difference between the electronic structure of the close-packed metals on the one hand and the body-centred metals on the

---

other. For the face-centred metals we have little to add to previous conclusions. These are to be thought of as arrays of close-packed ions or atoms having the configuration of full or nearly full 3d shells. The ' holes ' in the d-shells move from atom to atom and form a Fermi gas which is responsible for the high specific heat and ferromagnetism (paramagnetism in palladium). In each metal and in many of their alloys there are about 0·5 electrons per atom in a 4s or conduction band. The cohesion is partly due to these electrons, but at least as much to an attraction between the closed d-shells, probably of Van der Waals type. The Fermi surface extends over both the 4s and 3d bands, and transitions from the former to the latter caused by lattice vibrations have an important effect on the electrical resistance. The ferromagnetism is probably to be explained by Friedel's (1955) model, which depends on more than one magnetic carrier (hole) being able to penetrate on to one 3d shell.

In the body-centred metals there seems to be a band of which the first six states per atom are bonding and are occupied in chromium and iron ; the addition of further electrons to this band lowers the cohesion so there must be a minimum in the density of states for this band about here. This band may be formed mainly from the three 3d wave-functions of t_g symmetry (i.e. of type $xyf(r)$) overlapping a 4s band ; if so the radial part f of the wave-function has greater extension than for the free atom. There is also a narrow band formed mainly of wave functions of symmetry e_n (i.e. of type $(x^2-y^2)f(r)$). For chromium this probably lies above the Fermi surface and is unoccupied. In iron it contains exactly two electrons per atom. Following Griffith (1956) the magnetic moment of iron is ascribed mainly to these electrons and the remaining moment to a small magnetization of electrons at the surface of the bonding band in the sense described by Zener (1951) and Pratt (1957).

An important new conclusion is that these two magnetic electrons contribute nothing to the conductivity and do not in the normal sense form a band. A concept due to Mott (1949, 1952, 1956 a) is used, according to which an array of incomplete shells makes a sharp transition from a conducting or metallic state to a non-metallic or non-conducting state as the distance between them is increased. The magnetic electrons in iron are thought to be in these non-conducting states.

The consequences of this hypothesis both for the electrical conductivity and the x-ray absorption spectra of iron are outlined.

These conclusions are based on a study of the magnetic, electrical and other properties of these metals and also on the recent determination by Weiss (1957) of absolute x-ray scattering factors.

§ 2. CONDUCTING AND NON-CONDUCTING WAVE-FUNCTIONS

Pure crystalline substances at low temperatures are either good conductors or good insulators. An explanation of the sharp difference

was first given by Wilson (1931), using the model in which each electron
is described by a wave-function

$$\psi_k = u_k(x, y, z) \exp(i\mathbf{kr}), \quad \ldots \ldots \quad (1)$$

u_k having the periodicity of the lattice. According to Wilson, insulators
are substances in which all the Brillouin zones in k-space are either full
or empty, conductors (metals) those in which at least one zone is partly
occupied. However, there are many substances, such as NiO, which are
insulators when pure but which, according to this classification, should
have an incompletely filled zone. In the case of NiO, the eight electrons
of the Ni^{2+} ion cannot fill the ten places in the five d-bands. The absence
of conductivity is not due to a distance between the nickel ions so great
as to preclude tunnelling ; if Ni^{3+} ions are present, as in specimens
containing Li^+ ions, a current passes readily enough.

It is not suggested that a description of NiO as a polar crystal is
adequate ; there must be considerable covalent binding, which means
overlap of the electrons from the oxide ions on to the nickel. But how-
ever we describe it, it is difficult to see how a description of NiO in terms
of wave functions of type (1) can lead to zones which are all either empty
or full.

The experimental evidence suggests, then, that a crystalline array of
atoms with incomplete shells should normally be a conductor if the
overlap between the incomplete shells is large, but will not if the overlap
though finite is insufficient and if at the same time there is the same
(integral) number of electrons on each atom. Mott (1949, 1952, 1956 a)
has given arguments for believing that this transition from a conducting
to a non-conducting state should occur at a critical value of the inter-
atomic distance. The most direct evidence for this is provided by
observations of the conductivity at very low temperatures of germanium
or silicon as a function of the concentrations of the impurity centres of
majority type. An example is shown in fig. 1. Each impurity centre
may be considered as an ' atom ' containing one electron (or hole).
The sharpness of the transition from a ' metallic ' state in which the
conductivity depends little on temperature to a state in which thermal
activation is necessary for conduction is striking.

It cannot yet be said that a fully satisfying explanation of this transition
exists. As Löwdin (1956) stresses, one has to show that, at some value
of the lattice constant, the correlation between electrons leads to a
disappearance of the conductivity. Slater's (1951) elegant treatment
in which an antiferromagnetic arrangement of spins introduces a super-
lattice potential and splits a zone into two sub-zones, one filled and the
other empty, cannot be the whole story, because antiferromagnetic
materials can remain non-conductors above the Néel point. Mott
(1956 a) has given reasons for believing that the proper treatment of
these non-conducting states is the following : making use of *all* the
functions of the type (1) in a band, Wannier–Löwdin functions can

be set up, one localized on each atom. Each is multiplied by its appropriate spin function, and a Slater determinant written down to describe the wave-function $\Psi_{nc}(q_1, q_2 \ldots)$ of the whole system. This function is real and the energy of the whole system which it describes is separated from that of any excited states by several electron volts ; it thus represents a non-conducting state, at low temperatures ferro- or

Fig. 1
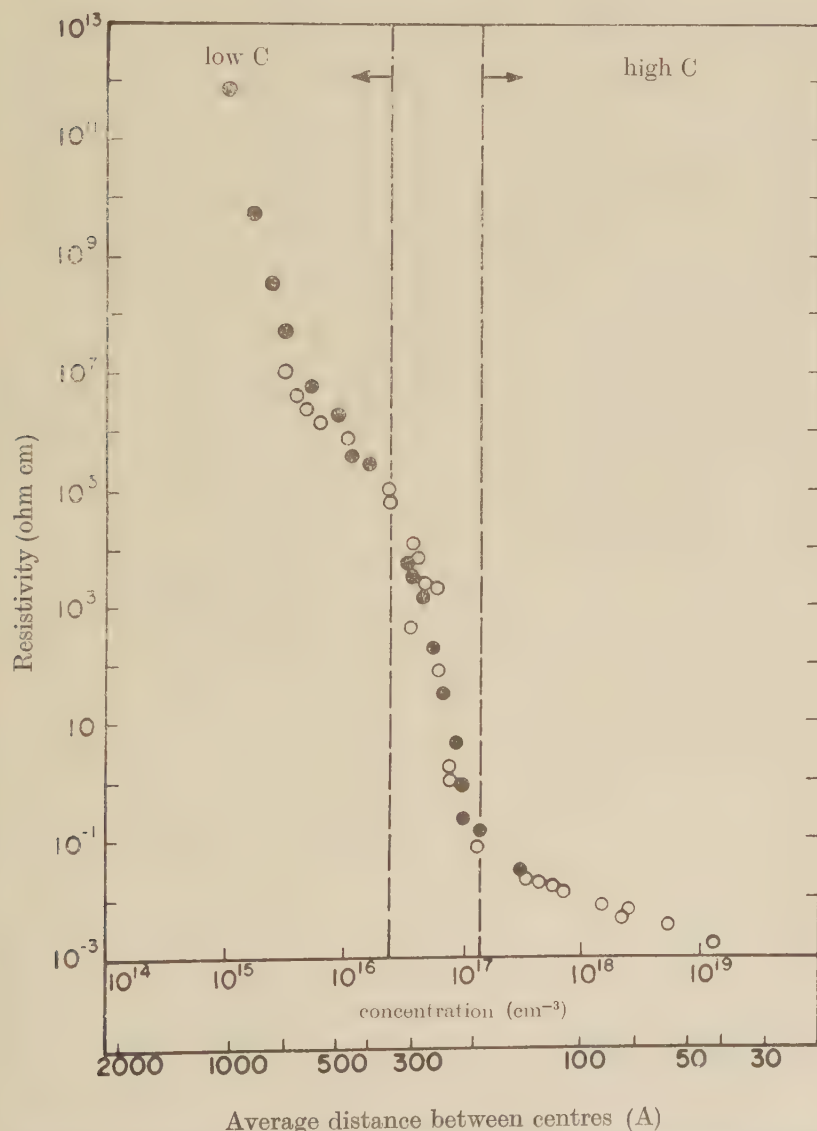
Electrical resistivity of germanium as a function of concentration of gallium at 2·5°K. In the range marked ' high C ' the conductivity between 20° and 2° is almost independant of T. Note the rapid drop at about 10^{17} cm^{-3} (Fritzsche 1956).

antiferromagnetic. It will be seen that ' correlation ' is included in this model in the zero order. When some parameter, such as the interatomic distance, is varied, the transition to a conducting state occurs at a definite value of this parameter. The conducting state of the whole system is described by a wave-function $\Psi_c(q_1, q_2 \ldots)$, which to a first approximation is a Slater determinant formed from *some only* of the wave-functions ψ_k, namely those describing the ' occupied states '.

We shall use in this paper the concept that an assembly of atoms or ions in atomic states each having a magnetic moment may be non-conducting (as the 3d electrons in NiO) or conducting (as the 3s electrons in sodium). A new idea of this paper is that in certain metals (certainly the rare earths, we believe also iron), the magnetic electrons *must* be described by these non-conducting wave-functions, while the conduction or bonding electrons are to be described by conducting wave-functions. We believe that this is the proper distinction to make, rather than that between bonding and localized or non-bonding, because there is no reason why non-conducting wave-functions should not contribute to the cohesion, and there may well be considerable overlap between the atomic or Wannier orbitals.

In former papers on non-conducting wave-functions (Mott 1949, 1956 a), it has been emphasized that, as the interatomic distance is decreased, there must be a sharp transition from a state where there are no free electrons to a state where there is a *finite* number. The argument is as follows. Two carriers of opposite sign, if not screened by other free carriers, attract each other with a Coulomb potential energy $-e^2/\kappa r$. Their lowest state will therefore be one in which they are bound to each other (as in an exciton). If other carriers are present, the Coulomb form must be replaced by a screened potential energy $(-e^2/\kappa r) \exp(-qr)$. If q is large enough, no bound state can be formed. Therefore, at $T=0$, for a conducting state there must be enough free carriers to produce this screening.

The rapidity of the drop in the resistance shown in fig. 1 suggests strongly that the transition from non-conducting to conducting states does involve a discontinuous change in the number of free electrons. It will be remembered that we have to deal here with a random arrangement of centres ; one may assume that for a crystalline array the transition would be even sharper.

In discussing the inner electrons in a *metal*, however, there is no theoretical argument to suggest that, as the interatomic distance is changed, there must be no carriers or else a *finite* (large) number. The conduction electrons will always lead to a screened potential between carriers in an inner band. But it is none the less legitimate to say that as the interatomic distance increases we must come to a state where there are no free carriers. As the interatomic distance increases the effective mass increases, and if the effective mass is large enough bound states are formed even with a screened potential.

§ 3. Conducting Wave-Functions and the Fermi Surface

A characteristic property of conducting wave-functions is the Fermi surface. This appears first in the theory as a consequence of the one-electron model ; a Slater determinant is set up from functions ψ_k of which the energy $E(\mathbf{k})$ lies within the (Fermi) surface of constant energy containing a number of states equal to the number of electrons. Any attempt to improve the wave-function Ψ_c will make use of the wave-functions of excited states (configurational interaction). Nevertheless the Fermi surface is not simply a consequence of a simple one-electron model : it describes a set of quantities which can be determined experimentally. This has been emphasized recently by the present author (Mott 1956 b), and various authors have studied the many-body problem in this connection (Gell-Mann 1957, Hubbard 1957, Nozières and Pines 1957). Here we need only mention the sharpness of the upper limit of x-ray emission bands due to transitions from the conduction band to an x-ray level (fig. 2) and their coincidence with the corresponding absorption edges in metals. Also we would stress the fact that experimental methods exist of determining the form in k-space of the Fermi surface : it is thus a question with physical significance to ask whether the Fermi surface lies partly in a zone occupied by the magnetic electrons. In nickel, palladium and cobalt the answer is, we believe, yes ; in iron and the rare earths, no.

Fig. 2

(a)   (b)
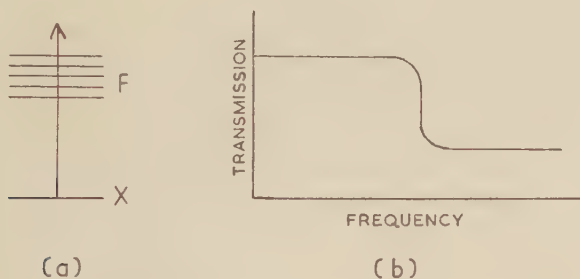
Level scheme (*a*) and absorption spectrum (*b*) for normal metal. X is the x-ray level and F the occupied levels in the conduction band.

To see this, let us consider a rare earth metal. It is reasonable to suppose that each atom has an integral number of electrons in each 4f shell and that these can be described by atomic orbitals, or, if there is some overlap between these atomic orbitals, by wave-functions Ψ_{nc} for the whole system which are of the non-conducting type. The points to emphasize are that these electrons contribute nothing to the current or to the Fermi surface, though, if an *extra* electron were added to one of the 4f states, it could jump from atom to atom, its energy forming a band *above* the Fermi surface. A method of demonstrating this would be to investigate the emission and absorption spectra of these metals

due to transitions with an inner d level (so that transitions to and from an f level are allowed). The type of spectrum to be expected† and a level scheme are shown in fig. 3. The point to emphasize is that the occupied non-conducting inner 4f levels are below the Fermi surface, and an empty band of levels is above it. The ' white line ' (maximum) in the absorption, common for many transition metals, will therefore be displaced from the edge, as in fig. 3 (b)‡.

Fig. 3

(a)   (b)
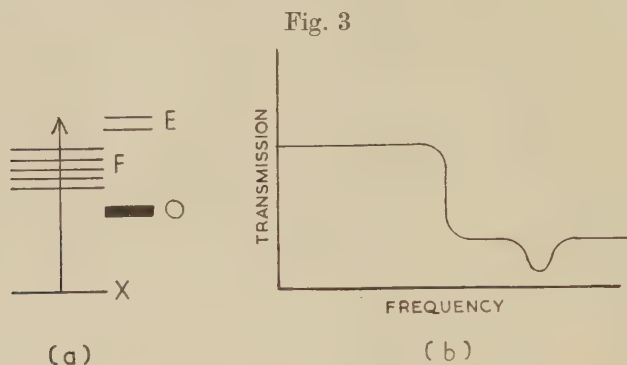
Level scheme (a) and anticipated x-ray absorption spectrum for rare earth metals (initial state d) or iron (initial state p). X and F as in fig. 2, O occupied inner levels, E empty inner band.

The width of the line due to the transition nd → 4f (n < 4) will be determined by two factors.

(a) The width ΔE of the 4f band. Although we have stated that the 4f electrons should be described by localized Wannier or atomic functions rather than by functions ψ_k of type (1), yet as soon as a hole is introduced into one of the occupied states, this hole can jump from atom to atom ; such a hole will therefore have a band of energy levels determined by the hopping integral. In the same way the empty levels above the Fermi level will form a band of the normal type for an extra electron there.

(b) The lifetime of the hole or electron. This will be determined by Auger interaction with the conduction electrons. The approximation in which we separate inner electrons into conducting and non-conducting

---

† This argument has been criticised by Professor Harvey Brooks (private communication). The basis of the criticism is that the x-ray absorption process leaves a positive charge on an inner shell, and that this charge, although screened by the conduction electrons, *may* give rise to a bound level below that of the band in which the extra electron moves. Moreover this bound level may be so lowered that it is no longer above the Fermi surface. The criticism is justified : if, however, a strong absorption line is observed above the absorption edge in any given case, we can only deduce that the screening is great enough to make the effect of the positive charge small, so that if the bound state exists at all, it is not pulled down to the Fermi level. The evidence that we shall present in § 10 suggests that this is so in iron.

‡ The suggestion that the maximum in the absorption ought to be displaced from the edge in iron was made by Dr. Orgel (see Griffith and Orgel 1957).

states is only a useful approximation if this lifetime is reasonably long, so that the maximum in absorption shown in fig. 3 exists.

It may of course happen (accidentally) that one of the levels shown in fig. 3 lies within an energy ΔE from the Fermi surface. In this case electrons from the Fermi band will move into these states and there will be a *non-integral* number of electrons in the levels under consideration. This is essentially the case in nickel, where the 3d level contains about 9·5 electrons : in other words nickel contains a mixture of ions with the configurations $3d^9$ and $3d^{10}$. We shall suggest that a similar situation exists in cobalt, but not in body-centred iron or chromium. Under such conditions the inner shell contributes both to the Fermi surface and to the conductivity, and the x-ray emission and absorption spectra will be as in fig. 4. Some observed spectra are shown in fig. 9.

Fig. 4

(a)   (b)
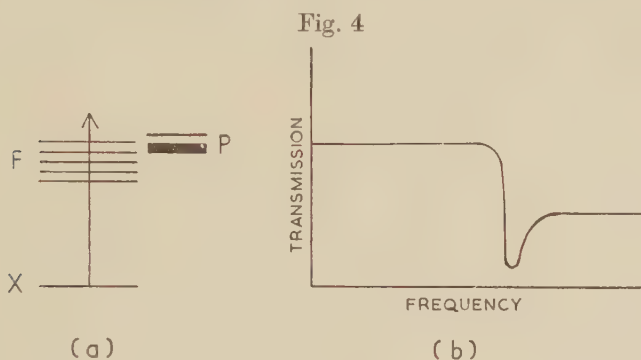
Level scheme (*a*) and anticipated absorption spectrum (*b*) for nickel and cobalt.

Thus a set of electrons can contribute to the Fermi surface and to a current for either of two reasons.

(i) If the number of electrons per atom is integral, it can only contribute when the overlap between atomic wave-functions is so large that the wave-function is conducting (as in sodium or magnesium).

(ii) Because the band (in the sense of (*a*) above) overlaps the Fermi surface of another set of electrons, so that the band contains a non-integral number of electrons. This is the case for the d-band of nickel or cobalt, where it is certainly proper to describe electrons or holes as jumping from atom to atom and contributing to the current and Fermi surface.

§ 4. MAGNETIC ELECTRONS IN IRON

Griffith (1956) suggested that the magnetic electrons in iron are in e_g states, which means that they are to be described by 3d wave-functions with symmetry x^2-y^2 etc., of which there are two per atom with each spin direction. Stevens (at the conference) pointed out that, in any attempt to use the L.C.A.O. method with hybridized orbitals to

set up Hartree–Fock wave-functions for iron, the radial functions $f(r)$ of the $x^2 - y^2$ functions need not be taken to be the same as those of the xy functions, as they are in Slater and Koster's paper (1954), and suggested that they ought to be smaller. Mott proposed (at the conference), following a suggestion by Longuet-Higgins, that the electrons in these e_g states ought to be described by 'non-conducting' wave-functions as for the rare earths rather than by Bloch wave-functions of type (1). The suggestion then is that these e_g electrons, while they may well overlap enough to give an exchange integral comparable with the Curie temperature, none the less are in non-conducting states and contribute nothing to the Fermi surface; in these states there is an *integral* number of electrons per atom. The three t_g functions (with symmetry xy), possibly much hybridized with 4s and 4p and perhaps with radii greatly different from in the atomic states, form a set of wave-functions of the type ψ_k which are strongly bonding. These states have room for six electrons, but overlap another band and so are partly filled, part of the Fermi surface falling within this zone. We call these electrons the bonding or conduction electrons.

Two electrons per atom are in the non-conducting states for iron. For chromium the corresponding energy levels are above the Fermi surface, so they form an empty, probably narrow, band of the normal type†. The states in iron are considerably lowered by the exchange forces between two electrons in the same atom with parallel spins; above the Curie point the two spins in a given atom remain coupled, as is shown by the finding of Hofmann *et al.* (1956) that the entropy per atom in iron due to the whole ferro-paramagnetic transition is equal to $k \ln 3$. It may be objected that at low temperatures, when the magnetism is saturated, the Slater determinants formed from Bloch functions of type (1) are the same as those formed by the Wannier functions, since the zone is full with two electrons per atom. With this we agree; with our model, however, at moderate and high temperatures when the magnetism is no longer saturated there is still no contribution to the Fermi surface or to the conduction from these electrons, as there would be with Bloch functions.

The magnetic moment of iron is 2·2 Bohr magnetons per atom, and the excess 0·2 over the value for two electrons is, we believe, due to orbital motion and to a small magnetization of the conduction electrons by the magnetic electrons, of the type discussed earlier by Zener (1951) and Pratt (1957). This will be discussed below.

We shall now discuss the evidence for this description of iron.

---

† Lomer has suggested that this band may overlap the Fermi surface and contain a small non-integral number of electrons which could be responsible for the observed antiferromagnetism; a similar suggestion has been made by Griffith and Orgel (1957), who cite recent observations of the x-ray absorption edge (Agarwal and Givens 1957) in support. The difficulty of this hypothesis is that it gives no explanation of the difference between this metal and manganese, and in particular of the low melting point of the latter.

§ 5. Results of Weiss on the Density of Electrons

By measuring the absolute value of the x-ray scattering factors, Weiss† (1957) has recently established a very interesting difference between the close-packed metals nickel and cobalt on the one hand and the body-centred metals iron and chromium on the other. While the intensities show that in the former metals most of the electrons are in orbitals similar to d orbitals, in iron there are only about two and in chromium perhaps none. The number of d electrons deduced by Weiss from his measurements is as follows :

Metal	Number of d electrons per atom	
Ni	9·5 ± 0·3	} type A
Co	8·5 ± 0·3	
Fe	2·3 ± 0·3	} type B
Cr	0·2 ± 0·4	

One concludes from this that in nickel and cobalt the 3d shells are relatively little distorted from the atomic form and that the cohesion is mainly due to a small number (less or equal to one per atom) of outer (conduction) electrons, in s+p orbitals ; the nature of this cohesion will be discussed in § 11. On the other hand, for the body-centred metals it is reasonable to suppose that for iron two electrons are in the non-conducting orbitals that we have described with radii similar to those in the free atom, while the remaining six electrons for iron and chromium form a partially filled band giving a distribution of electrons fairly uniform in space ; in the sense of the tight-binding approximation these functions will be formed of d (symmetry xy), 4s and 4p and higher functions.

§ 6. The Bonding Electrons in Iron and Chromium

We have now to discuss why two rather sharply separated forms of bonding should occur ; the iron and chromium type B in which six electrons seem to be responsible for the binding forces and for conduction, pushing any additional electrons into non-conducting orbitals probably mainly of e_g symmetry, and the cobalt nickel type A in which most of the electrons remain in unperturbed d-orbitals and in which the d- and s-bands both contribute to the Fermi surface.

There was, at the conference on which much of this paper is based, much discussion on whether types A and B were *necessarily* associated with the two crystal structures. The suggestion was made by Lomer that they were not necessarily correlated ; the electronic structure A

---

† We wish to thank Dr. Weiss for showing us his results prior to publication.

tends to have lower energy than type B if there are many more than six electrons in the bonding band of type B or if electrons have to be put into the (atomic) e_g orbitals which are not stabilized by exchange energy between parallel spins. Alloys of type A normally have a close-packed structure because this is the arrangement for which attracting spheres have the least energy. Alloys of type B normally have the body-centred structure for some reason unexplained, but it is suggested that in principle either crystal could occur with either electronic arrangement. It was thought that gamma-iron might have the same electronic arrangement as alpha-iron, because (Powell 1953) there is very little (about 0·3%) change in the electrical resistance at the alpha–gamma transition; on the other hand the face-centred iron–manganese alloys

Fig. 5
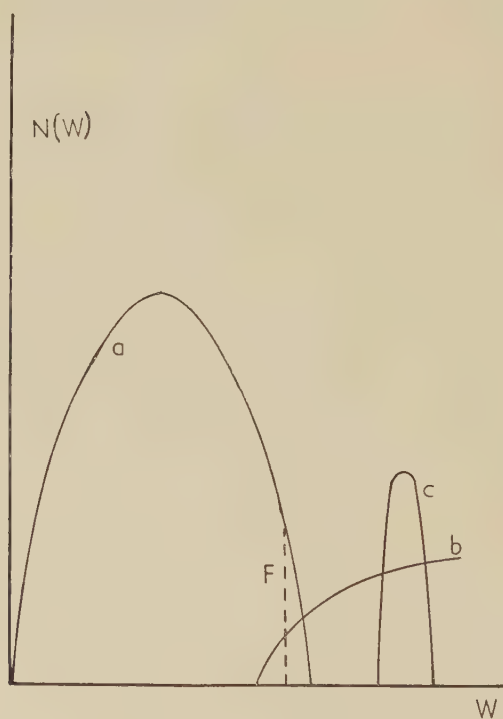
Suggested density of states for iron and chromium; the occupied states of symmetry e_g in iron are not shown. (*a*) is a band formed of hybridized 3d orbitals of symmetry t_g; (*b*) is an antibonding band which overlaps it; (*c*) are the empty e_g states; F is the Fermi surface.

(more than 12% manganese) are antiferromagnetic and were shown by Weiss from data on neutron diffraction to have moments much less than $2\mu_B$ and therefore almost certainly to have type A electronic

arrangement. Evidence was also given to suggest that there is a transition from type B to A as the concentration of cobalt is increased in the body-centred iron–cobalt alloys (cf. § 8).

For type B bonding, then, whatever the crystal structure, we assume a narrow band for the e_g states, which leads to non-conducting wave-functions if it contains an integral number of electrons. We consider that the six 3d functions with symmetry xy etc., to some extent hybridized with 4s (Altmann *et al.* 1957), form a band of which at any rate the lower states are highly bonding and contain six electrons. We think their bonding nature greatly extends in space the radial functions of these 3d states (i.e. hybridizes them with 4d), and that this is the reason why they do not contribute to the *F*-factor in Weiss's work. They are extended because in both body- and face-centred structures these xy d functions have large amplitudes in the direction of nearest neighbours. To these electrons (and not to the e_g electrons in non-conducting orbitals) must be ascribed the large linear electronic specific heat of iron[†]. Although six electrons would give a value of γ of this order with normal effective mass, the low electrical conductivity compared with copper rules out such a state of affairs. The Fermi surface must cover a fairly small area in k-space with a large effective mass. Perhaps, therefore, a band formed from the d_{xy} states just overlaps a 4s band. What is essential to the model is that the density of states a little above the Fermi surface is low, so that the addition of a few more electrons to this band rapidly increases the energy (decreases the binding). Thus when we go from iron to cobalt, since the additional electron cannot go into an e_g orbit without setting itself antiparallel to the other two and so having high energy. it must go into these anti-bonding states ; at this point the energy is lowered if the type A arrangement of electrons is chosen. The hypothesis also accounts for some of the properties of manganese ; on passing from chromium to manganese, the extra electron must go, either into the e_g orbital, where its energy is not as low as for iron because of the absence of exchange energy between coupled spins, or into the anti-bonding states. The evidence suggests that it is shared between the two ; some occupation of the anti-bonding states will account for the low cohesive force[‡] ; a *partially filled* e_g state will account

[†] Hofmann *et al.* (1956) give the following values for nickel, iron and gadolinium, which we shall use in this paper ; values for copper and chromium from Horowitz and Daunt (1953) are also included.

	Cu	Cr	Fe	Ni	Gd
$\gamma \times 10^4$	2	3·8	12·5	12	16

The results of Hofmann *et al.* are deduced by comparing specific heats at very low and high temperatures.

[‡] The melting points, a suitable measure of cohesive force, of metals near to manganese are, in degrees c (Hume-Rothery and Coles 1954)

V	Cr	Mn	Fe	Co	Ni
1700	1950	1250	1500	1450	1450

for the high electronic specific heat ($\gamma \sim 40$), and (Orgel 1957) for a complicated crystal structure through the Jahn–Teller effect.

We must now compare our conclusions with the L.C.A.O. calculations of Slater and Koster (1954) making use of earlier work by Fletcher and

Fig. 6
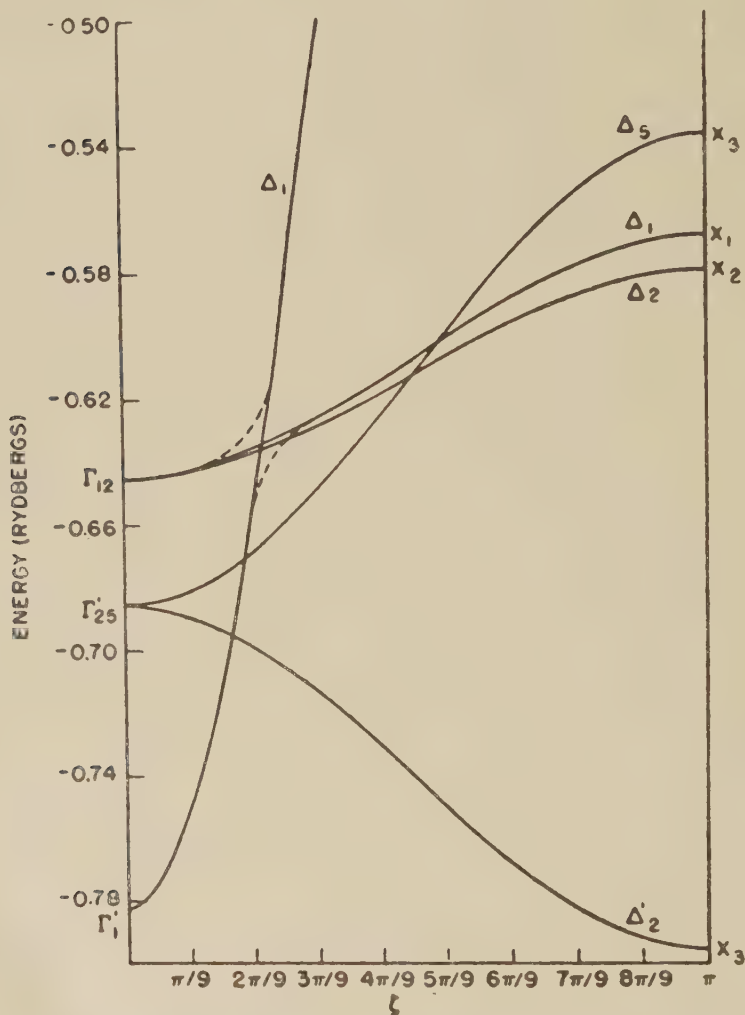
Energy bands in copper (Slater and Koster 1954). The curve shows the variation of energy with wave vector in the (001) direction.

Wohlfarth (1951), and more recent work using the method of orthogonal plane waves (Callaway 1955). In Slater and Koster's fig. 1 (reproduced in our fig. 6) one sees among the five branches of the d band two which are comparatively narrow; one may assume that, though all are hybridized from xy and x^2-y^2 functions, the narrow ones are predominantly x^2-y^2. One might anticipate a similar result if calculations

were made for the body-centred structure. For this structure our assumptions differ from theirs in the following respects.

(*a*) It is legitimate to take different radial functions for the broad and narrow branches, those for the broad bands, to account for Weiss's results, being extended considerably in space (hybridized with 4d, probably also with 4s).

(*b*) Correlation in the narrow bands is so strong as to make them ' non-conducting ' in the sense explained in this paper.

(*c*) The empty half of this band in iron and the whole of it in chromium are above the Fermi surface. As Slater and Koster have criticized the assumption made by Jones and Mott (1937) that the band mainly formed from the x^2-y^2 functions might be separated from that mainly formed from the xy (they are not, according to their calculations), we must point out that with somewhat directional orbitals formed from d-functions we may expect a non-spherical form of the potential, which will split the 3d *states*. We do not see how it can be said *a priori* that the (mainly) xy and x^2-y^2 bands should not be separated in a given case.

§ 7. The Magnetic Moment of Iron

Our hypothesis is that non-conducting e_g states contain two electrons per atom, which contribute a moment of two Bohr magnetons per atom both above and below the Curie point (compare Hofmann *et al.* 1956). The observed saturation moment is $2 \cdot 2\,\mu_B$. Some of the differences will be due to orbital motion. Some, however, may be due to the magnetization of the electrons near the Fermi level by interaction with the magnetic electrons, as first suggested by Zener (1951). A crude estimate of its magnitude may be made as follows. An electron in the conduction (bonding) band will see different potentials according as its spin is parallel or antiparallel to the magnetic electrons. Suppose its energy ($\int \psi^* H \psi\, d\tau$ over an atomic cell with ψ normalized to unity in this cell) is lower by $\varDelta W$ in the atomic cell in which it finds itself. Then if x excess conduction electrons per atom have spin parallel to the direction of magnetization, the energy is increased, compared with the state without magnetization, by

$$-\tfrac{1}{2}x\ \varDelta W + \tfrac{1}{2}x^2/N(W),$$

where $N(W)$ is the density of states per atom at or near the surface of the Fermi distribution. This is a minimum when

$$x = \tfrac{1}{2}N(W)\ \varDelta W.$$

To agree with the observed magnetic moment, subtracting a reasonable contribution for orbital motion, we set $x \sim 0 \cdot 1$. The density of states may be obtained from the observed electronic specific heat ; we find $N(W) = 1$ (ev)$^{-1}$. We deduce $\varDelta W \sim 0 \cdot 2$ ev. A recent paper by Pratt (1957) giving calculations along these lines suggests that such a value is reasonable.

On the other hand the considerations in § 10 suggest a rather stronger interaction between the conduction electrons and the magnetic electrons than this. A strong interaction is perhaps not incompatible with a small magnetization of the conduction electrons, for the following reasons. It will in general lower the energy of the conduction electrons to hybridize them slightly with terms of the e_g type, especially for the body-centred structure where the x^2-y^2 orbitals point towards next nearest neighbours ; to include such terms will only be possible when the spin is antiparallel to that of the magnetic inner electrons, and will tend to magnetize them in this sense. Such an effect may largely cancel out a magnetization of the type described above.

§ 8. Alloys of Iron with the Body-Centred Structure

Hume-Rothery and Coles (1954, p. 161) remark that " there is a marked difference between the ferromagnetism of iron on the one hand and of cobalt and nickel on the other. When elements are dissolved in alpha-iron to form substitutional solid solutions, the saturation moment is generally reduced by $2 \cdot 22\,\mu_B$ for each atomic substitution. . . . In contrast to this, when elements of well-defined valency (e.g. Cu, Zn, Al) are dissolved in nickel, the saturation moment is reduced by $v\,\mu_B$ per atomic substitution, where v is the valency of the solute ". This is clearly in agreement with our model† and gives strong support to it ; the partly filled d-band of nickel can accept extra electrons as described in Mott and Jones (1936, p. 189), but the non-conducting states of iron cannot ; here electrons can be added or subtracted from the surface of the Fermi distribution, which will not greatly affect the magnetism except for the residual term $0 \cdot 2\,\mu_B$.

Exceptions to this rule are shown by the elements immediately adjacent to iron. The addition of nickel or cobalt gives moments suggesting that these atoms make contributions to the magnetism of rather more than $2\,\mu_B$. We suggest that, in dilute solutions, these atoms take two electrons with coupled parallel spins into x^2-y^2 orbitals and set their spins parallel to those in the surrounding iron ; the third or fourth electrons cannot go into these orbitals because they would not gain exchange energy from having spins parallel with the others in the same atom, since the 3d states with symmetry x^2-y^2 can only take two electrons with a given spin. If the extra electrons, one for cobalt and two for nickel, were added to the surface of the Fermi distribution, the alloy should have the same

---

† An alternative explanation (Stoner 1946, Hume-Rothery and Coles 1954) is, in terms of the usual band model, that in iron the magnetic carriers in a band are not fully oriented. While there is nothing objectionable about this hypothesis, which indeed is probably valid for face-centred iron alloys, we doubt if it would lead to the observed results. It seems unlikely that electrons added to a d-band on alloying would be shared equally between the two spin directions, since the density of states will have different values for the two directions.

moment per atom as iron ($2 \cdot 2\,\mu_B$). The observed behaviour is as follows (compare Coles and Bitler 1956). At first the addition of cobalt gives *three* Bohr magnetons per atom of cobalt (see Stoner 1946). We have suggested that the two electrons go into non-conducting 3d orbitals of symmetry $x^2 - y^2$. When this is done, the cobalt atom has one excess electron, which must be used for screening its excess nuclear charge. We suggest that as described by Friedel (1954) a *bound* state is formed, from wave-functions in the conduction band (i.e. 3d functions with symmetry xy), and that the electron in this state will set itself parallel to the other two with symmetry $x^2 - y^2$.

Friedel (at the conference) expressed doubt as to whether a bound state would be formed due to a single excess charge, for so high a value of $N(W)$. This may perhaps occur if there is a *strong* exchange interaction between the electron in the bound state and the two e_g electrons, since this would lower the energy of the supposed bound state. We have seen that this does not necessarily imply a strong magnetization of the conduction electrons.

When, however, the concentration of cobalt passes about 35% the magnetic moment begins to drop again : in the similar case of FeNi the neutron diffraction evidence of Shull (1955) shows that *iron* carries a spin of $2 \cdot 6\,\mu_B$. The interesting suggestion was made by Lomer that somewhere in this series there is a transition to an electronic arrangement of type A, although the body-centred structure is retained ; an arrangement of type A, with the usual 0·5 4s electrons per atom, would imply a mean moment of about $2 \cdot 5\,\mu_B$ on each atom. Evidence in support of this idea was provided by Coles, who pointed out, while the addition of aluminium to iron did not decrease the moment per iron atom, its addition to an iron cobalt alloy did. The relevant experimental material is discussed by Coles and Bitler (1956), though the interpretation given there is in terms of an unsaturated d-band in iron.

§ 9. Electrical Resistance of Ferromagnetic Metals

One of the present authors (Mott 1935, Mott and Jones 1936) has given a theory of the electrical behaviour of nickel, palladium and their alloys which explains in particular the anomaly in the resistance of nickel at the Curie point. This theory assumes that the abnormal behaviour of these metals is due to transitions in which electrons, under the influence of the lattice vibrations, make transitions from the s band (which carries most of the current) to the d band. The consequences of this theory have been tested by Conybeare (1937), Allison and Pugh (1956), Overhauser and Schindler (1957) and others, and for these metals s–d transitions almost certainly occur, are responsible for the comparatively high resistivity and at any rate a large part of the resistance anomaly. But the rare earths and iron also show an anomaly at the Curie point ; if our theory is correct, this cannot be due to the same cause. The anomalies are shown in fig. 7, and the form suggests that here we have

to do with the scattering of conduction electrons by the disordered spins of the magnetic electrons, behaving above the Curie point like a disordered alloy ; below the Curie point a description in terms of spin waves would be preferable.   At the conference a description of such a theory was given by de Gennes (see de Gennes and Friedel 1957, Kasuya 1956).

Fig. 7
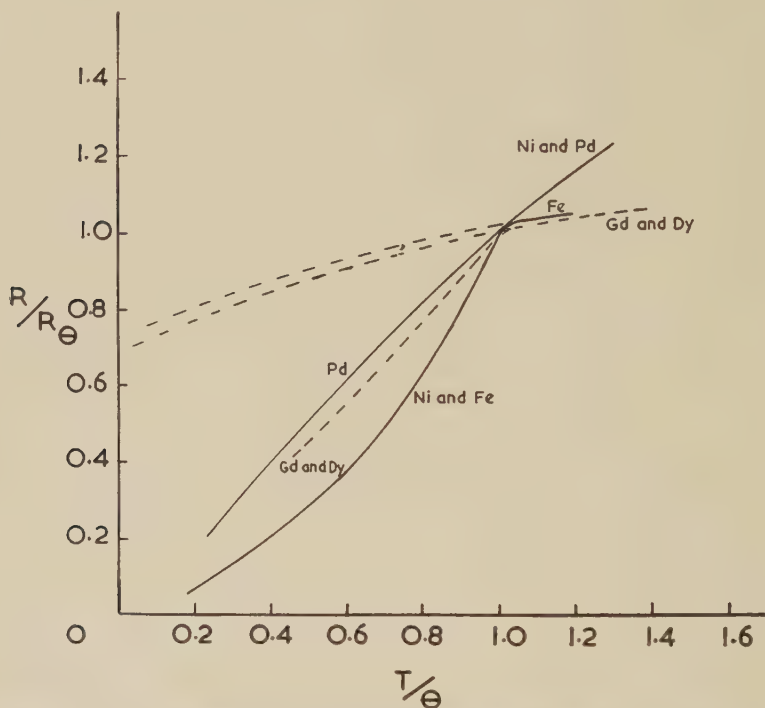
Reduced electrical resistance of some ferro- and paramagnetic metals ; Θ is the Curie temperature and R_Θ the resistivity there.

Ni and Pd	Gerritsen (1956)
Fe	Powell (1939)
Gd and Dy	Legvold (1953)

Note the similar behaviour above the Curie point of Ni and Pd and of Fe and the rare earths.

This would mean that in iron the extra resistance near and above the Curie point would *not* be due to lattice vibration.   A test would be to compare the resistance of FeRu with NiPd.   They should differ as shown in fig. 8†.

If the resistance of iron above the Curie point is due to disordered spins, we can estimate it very roughly as follows.   We have already seen that a conduction electron sees a potential energy differing by ΔW according as the electron has spin parallel or antiparallel to the magnetic

─────────

† A paper by B. R. Coles giving the results of an experimental investigation of this point will appear shortly in this journal.

electrons. If we knew anything about the wave-functions of the con-
duction electrons near the Fermi surface we could thus estimate the
resistance by the methods used by Mott and Jones (1936, p. 291) for
dilute alloys ; unfortunately we do not. However, a value of ΔW of
0·02 ev in a disordered alloy of the AgAu type would give only one-
fortieth of the normal resistance of these metals at 1000°K ; both may be
expected to be scaled up for iron to roughly the same extent, suggesting
that ΔW should be about 1·5 ev to account for the observed anomaly.

Fig. 8

(a)   T (ABSOLUTE)

(b)
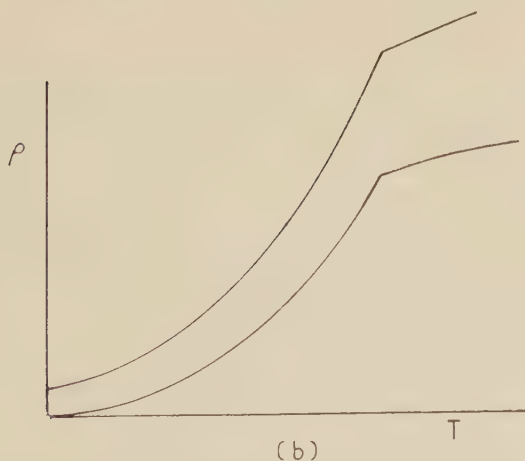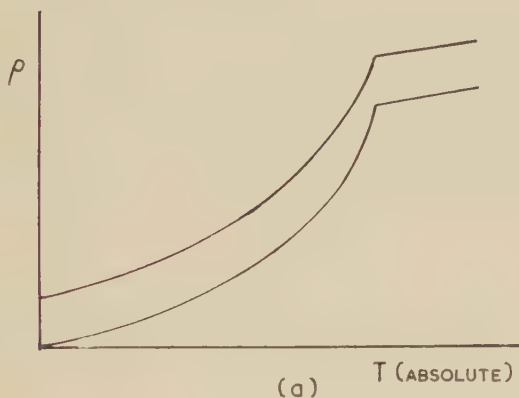
Predicted behaviour of the electrical resistance of iron and nickel alloys.
(*a*) Iron and iron–ruthenium ; (*b*) nickel and nickel-palladium.

It seems to us, then, that to account for the resistance anomaly at
the Curie point a fairly strong interaction must be assumed, and one
that would give a much larger magnetization of the conduction electrons

than the observed value of less than 0·2, unless there is some compensating effect as suggested in § 8. However, more detailed calculations along the lines of those of Pratt (1957) and Yosida (1957) are required.

§ 10. X-Ray Absorption

As already stated a test of our theory for iron would be to measure the L_3 or M_3 absorption edge. Measurements have, however, been made of the M edge by Carter and Givens (1956); these certainly suggest that the ' white line ' is displaced from the edge, though the evidence is not clear, being complicated by the doublet nature of the 3p level. Their results have, however, been analysed by Tomboulian (1957), and the result shows a displacement of the white line (maximum in the absorption) from the edge, just of the type predicted in fig. 4. These results, together with the original observations, are reproduced in fig. 9 (*a*). Figure 9 (*b*) shows, by way of contrast, some recent observations of the L_3 absorption of a nickel foil by Cauchois and Bonnelle[†] (to be published); here the maximum is at the edge, as predicted.

§ 11. Cohesion in the Face-Centred Transition Metals

Calculations of the cohesive force in copper, silver and gold have recently been made by Kambe (1955), assuming that it is due to the electrons in the half-full conduction band. The values obtained are (k cal per atom)

	Cu	Ag	Au
Calculated	59	56	49
Observed	81	68	92

It is generally believed that the difference between the observed and calculated values (about 1 ev for copper and silver) is due to some attraction between the filled d shells. The simplest assumption is that the wave-functions in the d-band are hybridized with 4s and higher terms—which must in any case be so, and that this hybridization, while small enough to be comparable with Weiss's experimental results, is big enough to give the observed cohesion. It seems to us more likely, however, that the greater part of this cohesion is of the Van der Waals type (Friedel 1953) (correlation between dipoles or quadrupoles) and cannot be described in the Hartree–Fock approximation. The cohesion must be due (e.g. in copper) to the proximity between the states $3d^{10}4s^1$ and $3d^94s^2$. According to Friedel (1953) the ion *in the solid* can make a transition from the state $3d^{10}$ to the state $3d^94s^1$ with an energy of about 2 ev. Owing to interaction with the conduction electrons such a transition has some dipole moment and is responsible for the colour of copper; in calculating the cohesion, however, it would be simpler
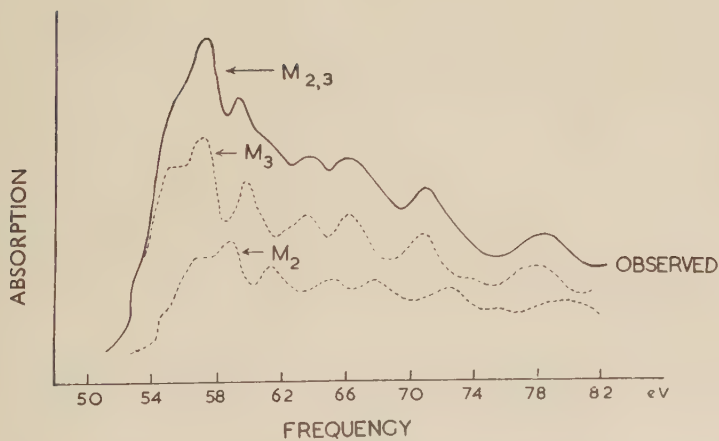
---

† We wish to thank Mlle Cauchois for allowing us to see these results and others for iron and cobalt prior to publication,

to neglect the conduction electrons and calculate the quadrupole inter-action of Van der Waals type. Without knowing the quadrupole moment of the transition, however, it is not possible to estimate the contribution to the cohesive energy.

Fig. 9
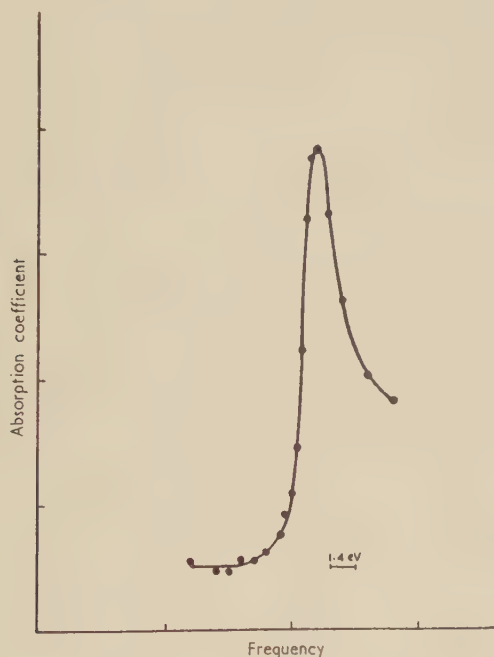
(a) Apparent L_3 absorption coefficient of nickel (Cauchois and Bonnelle, to be published). (b) M_2 and M_3 absorption coefficient of iron ; analysed by Tomboulian (1957) from observations of Carter and Givens (1956).

We turn now to the face-centred transition metals. Our model here is the 'collective' electron model of Mott (1935), Slater (1936), Stoner (1946) ; the discussion of Van Vleck (1951) is relevant, showing in what respects a simple collective treatment must be modified by correlation. In any case these materials have a nearly full d-band containing about 9·5 electrons in nickel and palladium, 8·5 in cobalt. Thus we may speak of 0·5 or 1·5 'holes' per atom and these are mobile, obey Fermi–Dirac statistics and contribute to the Fermi surface. Van Vleck's paper emphasizes that there may be strong correlation forces preventing more than one going on the same atom, but Friedel (1955) argues that these correlation forces are not great enough to prevent the presence on one atom of two carriers at the same time. He suggests that to exchange forces between two carriers *when they are on the same atom* we can ascribe the ferromagnetism†.

The absence of ferromagnetism in face-centred iron is of considerable interest. In spite of the very small difference in the electrical resistance both of iron and of iron–manganese alloys in the face- and body-centred structures (Powell 1947, 1953) above the Curie point of the latter‡, it seems reasonable to suppose that face-centred iron or iron alloys have the electronic structure of the other face-centred metals. Face-centred iron–manganese alloys are antiferromagnetic with small (about $0·5\,\mu_B$) magnetic moment per atom, which seems impossible for the type B structure. For the type A electronic structure, on the other hand, Friedel's theory gives us some indication of why ferromagnetism is absent. Face-centred iron will be a mixture of atoms in states $3d^7 4s^1$ and $3d^8$; the 'carriers' are pairs or triplets of holes with coupled parallel spins. If two triplets or a pair and a triplet try to get onto the same atom, this will lead to d-shells with four or five electrons only, which are not found in the atomic spectra. We conclude that there are strong correlation forces preventing carriers from getting on to the same atom, so that Friedel's mechanism cannot apply. That the gas of carriers shows some antiferromagnetism must perhaps be explained along the lines used by Lidiard (1954) and by Slater and Koster (1954) for chromium.

Finally we add a few remarks about cohesion in the face-centred transition metals. We have pointed out that in the noble metals about half the total cohesive energy comes from (Van der Waals) interaction between the full d-shells. In the transition metals the proportion is

---

† Such an explanation cannot be used for body-centred iron, if the electrons are in non-conducting states, or for gadolinium. The ferromagnetism may be due either to direct (Heisenberg) interaction between the d-shells if the exchange integral is ever of the right sign, or more likely to some coupling involving the conduction electrons, either super-exchange (Pratt 1957) or through a small magnetization of the conduction electrons as in Zener's (1951) theory.

‡ Below the Curie point for iron–manganese alloys there is a large difference (Powell 1947).

probably higher; in nickel, for instance, the total cohesive energy is higher than copper but 0·5 s electrons per atom can hardly produce as much cohesion as one. It is certain also that a nickel atom in the state $3d^{10}$ must be very polarizable.

Nevertheless the evidence remains very striking that a whole range of alloys of nickel and palladium with copper, zinc, gallium, silver, cobalt etc. appears to retain about 0·5 electrons per atom over a wide range of composition (cf. the discussions given by Mott and Jones 1936). The explanation must be along the following lines : suppose one calculated the energy of a nickel lattice formed of atoms in the state $3d^{10}$. Owing to Van der Waals attraction one would get fairly strong cohesion. Suppose one now removed n electrons per atom from some of the d-shells

Fig. 10
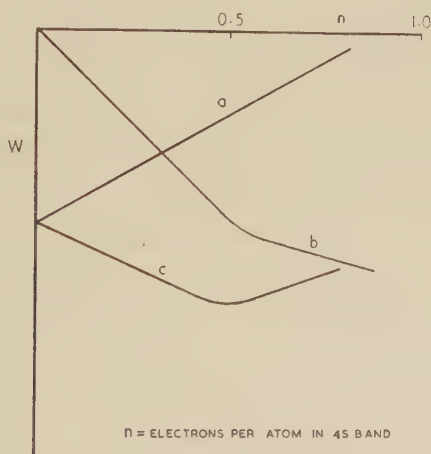
Suggested form of the cohesive energy W of nickel. (*a*) That due to Van der Waals attraction between atoms with $3d^{10}$ configuration ; (*b*) that due to the conduction electrons ; (*c*) total energy.

and put them into the conduction band. They would contribute negative energy of Wigner–Seitz type, the contribution per atom becoming numerically less with increasing n because of the Fermi energy (curve (*b*) in fig. 10). On the other hand the contribution to the cohesion from Van der Waals interaction would fall (curve (*a*)).

The experimental results show that for a wide range of alloys, and hence for a wide range of slopes of curve (*a*), the sum of the two (curve (*c*)) has its minimum at $n=0·5$. This must mean that curve (*b*) has a rather sharp change of slope for this value of n. Why this is so is not clear.

REFERENCES

AGARWAL, B. K., and GIVENS, M. K., 1957, *Phys. Rev.*, **107**, 62.
ALLISON, F. E., and PUGH, E. M., 1956, *Phys. Rev.*, **102**, 1281.
ALTMANN, S. L., COULSON, C. A., and HUME-ROTHERY, W., 1957, *Proc. roy. Soc.* A, **240**, 145.
CALLAWAY, J., 1955, *Phys. Rev.*, **99**, 500.

CARTER, D. E., and GIVENS, M. P., 1956, *Phys. Rev.*, **101**, 1469.

CAUCHOIS, Y., and BONNELLE, C., 1957, *Comptes Rendus* (in the press); *Phil. Mag.* (to be published).

COLES, B. R., and BITLER, W. R., 1956, *Phil. Mag.*, **1**, 477.

CONYBEARE, J. G. G., 1937, *Proc. phys. Soc. Lond.*, **49**, 29.

DE GENNES, P. G., and FRIEDEL, J., 1957, *J. phys. chem solids* (in the press).

FLETCHER, G. C., and WOHLFARTH, E. P., 1951, *Phil. Mag.*, **42**, 106.

FRIEDEL, J., 1952, *Proc. phys. Soc. Lond.*, B, **65**, 769 ; 1954, *Advanc. Phys.*, **3**, 446 ; 1955, *J. de Physique* (8), **16**, 829.

FRITZSCHE, H., 1956, *Purdue University Second Quarterly Report*.

GELL-MANN, M., 1957, *Phys. Rev.*, **106**, 369.

GERRITSEN, A. N., 1956, *Handbuch der Physik*, **19**, 187.

GRIFFITH, J. S., 1956, *J. Inorg. nuclear Chem.*, **3**, 15.

GRIFFITH, J. S., and ORGEL, L. E., 1957, *Nature, Lond.* (in the press).

HOFMANN, J. A., PASKIN, A., TAUER, K. J., and WEISS, R. J., 1956, *Phys. Chem. Solids*, **1**, 45.

HOROWITZ, M., and DAUNT, J. G., 1953, *Phys. Rev.*, **91**, 1099.

HUBBARD, J., 1957, *Proc. roy. Soc.* A. **240**, 539.

HUME-ROTHERY, W., and COLES, B. R., 1954, *Advance. Phys.*, **3**, 149.

JONES, H., and MOTT, N. F., 1937, *Proc. roy. Soc.* A, **162**, 49.

KAMBE, K., 1955, *Phys. Rev.*, **99**, 419.

KASUYA, T., 1956, *Prog. theor. Physics*, **16**, 45.

LANDAU, L. D., 1956, *J. exp. theor. Phys.*, **30**, 1058.

LEGVOLD, S., 1953, *Rev. mod. Phys.*, **25**, 129.

LIDIARD, A. B., 1954, *Rep. Progr. Phys.*, **17**, 201.

LÖWDIN, P. O., 1956, *Advanc. Phys.*, **5**, 1.

MOTT, N. F., 1935, *Proc. phys. Soc. Lond.*, **47**, 571 ; 1949, *Ibid.*, **62**, 416 ; 1952, *Advanc. metal Phys.*, **3**, 76 ; 1956 a, *Canad. J. Phys.*, **39**, 1356 ; 1956 b, *Nature, Lond.*, **178**, 1205.

MOTT, N. F., and JONES, H., 1936, *Theory of the Properties of Metals and Alloys* (Oxford : Clarendon Press).

NOZIÈRES, P., and PINES, D., 1957, *Phys. Rev.* (in the press).

ORGEL, L. E., 1957, *Nature, Lond.*, (in the press).

OVERHAUSER, A. W., and SCHINDLER, A. I., 1957 (*NRL Report* 4920).

POWELL, R. W., 1939, *Proc. phys. Soc. Lond.*, **51**, 416 ; 1947, *J. Iron St. Inst.*, No. II for 1946, 83P ; 1953, *Phil. Mag.*, **44**, 772.

PRATT, G. W., 1957, *Phys. Rev.*, **106**, 53.

SHULL, C., 1955, *Les Electrons dans les Metaux*, 10th Solvay Conference Report, 227.

SLATER, J. C., 1936, *Phys. Rev.*, **49**, 537 : 1951, *Ibid.*, **82**, 568.

SLATER, J. C., and KOSTER, G. F., 1954, *Phys. Rev.*, **94**, 1498.

STONER, E. C., 1946, *Rep. Progr. Phys.*, **11**, 43.

TOMBOULIAN, D. H., 1957, *Technical Report* ORD, 669.

WEISS, R. J., 1957, *Rev. mod. Phys.*, Cambridge (Mass.) Conference (to be published).

WILSON, A. H., 1931, *Proc. roy. Soc.* A, **133**, 450.

VAN VLECK, J. H., 1951, *Rev. mod. Phys.*, **17**, 27.

YOSIDA, K., 1957, *Phys. Rev.*, **106**, 893.

ZENER, C., 1951, *Phys. Rev.*, **81**, 440 : *Ibid.*, **83**, 299. See also Zener, C., and Heikes, R. R., 1953, *Rev. mod. Phys.*, **25**, 191.

CORRESPONDENCE

Quenching Vacancies in Dilute Binary Platinum Solid Solutions

By S. Pearson and F. J. Bradshaw
Royal Aircraft Establishment, Farnborough

In a previous paper (Bradshaw and Pearson 1956) we have described experiments in which changes in the resistance of platinum wires following quenching and annealing were interpreted in terms of vacancies. Since it is reasonable that vacancies might tend to associate with impurity atoms we have repeated such measurements on wires to which small amounts of impurity in solution were added. Platinum wires of 0·1 mm diam. containing (a) 1 atm% of Rh and (b) $\frac{1}{2}$ atm% of Au were obtained from Johnson Matthey & Co. for this purpose.

Theories and experiments on solute and self-diffusion in dilute solid solutions support the idea of association (Alfred and March 1956, Hoffmann et al. 1955, Lazarus 1955) though the results suggest that the difference between the energies to form a vacancy near to and far from a solute atom, i.e. the binding energy, is likely to be small compared with the vacancy formation energy. Measurements of the variation with temperature of the resistance increase on quenching the alloys confirmed this; the vacancy formation energies deduced for both alloys were within $\pm 0 \cdot 1$ ev of the figure for pure platinum, viz. 1·4 ev. Also the resistance increases on quenching the pure metal and the alloys from the same temperature were not significantly different.

The annealing behaviour of the alloys (quenched from $\sim 1400°$c) did not differ significantly from that of the pure metal but the results can be used to place a limit on the binding energies. If the concentration of vacancies is much less than the concentration of solute atoms one may deduce that the difference between the annealing activation energies in the alloy and the pure metal is $B [1 - \{1 + 12c \exp (B/kT)\}^{-1}]$ (where B is the binding energy and c the concentration of solute atoms). The measured activation energies were compared with that for pure platinum by the usual statistical procedure and it was found that there was a 95% chance that the difference lay between $0 \cdot 04 \pm 0 \cdot 07$ ev (Pt–Rh alloy) and $0 \cdot 11 \pm 0 \cdot 06$ ev (Pt–Au alloy). From the relation given we may deduce that the vacancy–solute atom binding energies would be very unlikely to exceed 0·16 ev (Pt–Rh) and 0·23 ev (Pt–Au). They would probably be much smaller than these values.

As the temperature of the alloy is lowered a binding between vacancies and solute atoms will cause the equilibrium proportion of free to bound vacancies to decrease. There will be a critical temperature during the quenching of an alloy below which the free/bound interchange of vacancies becomes too slow to allow equilibrium to be reached and a ratio of free to bound vacancies characteristic of this temperature will be retained. If the critical temperature is defined as being that at which the rate of fall of the equilibrium concentration of vacancies is equal to the rate at which vacancies become bound to the solute atoms then it may be calculated for a given binding energy. For the concentrations of solute atoms used here and binding energies between 0·01 and 0·3 ev it is in the range 450°–550°K. After quenching, the alloy specimen may be annealed at a temperature too low for the vacancies to disappear to sinks, but high enough for the free to bound interchange to proceed to equilibrium. If there is an appreciable binding energy between vacancies and solute atoms and their contributions to the resistivity are less when bound together than when they are separate then annealing at this temperature should lower the resistance of the specimen. Anneals at 85°C for several minutes showed that the fall in resistance, if occurring, was less than 1 : 250. This figure allows us to relate the binding energy with the fall in resistivity of a vacancy which occurs when it becomes bound. For a binding energy of 0·02 ev it can be shown that the fall in resistivity of a vacancy on being bound must have been less than 15% (Pt–Rh) or 27% (Pt–Au). For 0·05 ev it was less than 4% (Pt–Rh) or 6% (Pt–Au). For 0·1 ev it was less than 2% in both alloys.

In all the above we have considered annealing involving a single process only. If it were more complex and involved a combination of two processes, e.g. vacancies and divacancies migrating (Koehler, private communication), then the conclusions relating binding energy to fall in resistivity would in general apply to the process with the lower activation energy.

<div align="center">REFERENCES</div>

ALFRED, L. C. R., and MARCH, N. H., 1956, *Phys. Rev.*, **103**, 877.
BRADSHAW, F. J., and PEARSON, S., 1956, *Phil. Mag.*, **1**, 812.
HOFFMANN, R. E., TURNBULL, D., and HART, E. W., 1955, *Acta Met.*, **3**, 417.
LAZARUS, D., 1955, *Impurities and Imperfections* (Cleveland : A. S. M.), p. 107.

Energy Determination of Electromagnetic Cascades in Nuclear Emulsions

By K. Pinkau

H. H. Wills Physical Laboratory, University of Bristol

In an earlier publication (Pinkau 1956) it was shown that the primary energy E_0 of a γ-ray initiating a cascade shower could be estimated in practice by measuring the multiple scattering of the cascade electrons. Then, by comparing the results for the numbers of electrons and their energies with the cascade theory of longitudinal development under 'Approximation A' of Rossi and Greisen (1941), the primary energy could be deduced.

However, scattering measurements are very tedious and sometimes impossible owing to the steepness of the cascade or distortion of the emulsion. An attempt was made, therefore, to obtain sufficient information about E_0 merely by determining the radial distribution of cascade electrons at given distances from the origin (taken as the point where the first pair converts in the cascade).

The idea of the method can be seen from a consideration of Nishimura and Kamata's (1952) distribution function. Let $N(E_0, 0, R, t)$ be the total number of particles inside the disc with radius R after t cascade units ($=2\cdot9$ cm in emulsion) distance from the origin in a cascade with primary energy E_0. It can be shown that in an approximation for $R \ll 1$, N is a function of $E_0 \times R$ only (where R is measured in cascade units). A new variable will be defined here as $Z_0 = E_0 \times R$. Then,

$$N(E_0, 0, R, t) = N(E_0 \times R, t), \quad R \ll 1.$$

This means that for a given value of t *the distribution N is, for $R \ll 1$, the same for all primary energies E_0, and all materials, provided N is measured as a function of $Z_0 = E_0 \times R$.* The function N can be determined from observations on cascades of known energy, and hence the primary energy of any other cascade may then be deduced from the measurement of N vs R.

The primary energies of two cascades were determined by the method previously described (Pinkau 1956). Together with the cascades A, B and D of this reference they served as a set of five calibration cascades with values of E_0 given in the graph. The cascades of Pinkau (1956) were remeasured because no great care had been taken to include electrons below 100 MeV in the former experiment. Also, their values of E_0 have been slightly increased since a scattering constant of $28\cdot5$ should be used rather than the former value of 27 (Brisbout *et al.* 1956). Target diagrams were made at 2, 3, 4, 5·5 and 7 cascade units distance plotting all tracks which fulfilled the following conditions : (*a*) the grain density was

at plateau value, (*b*) the angles projected in the plane of the emulsion were <25° to the projected angle of the cascade core. A special eye-piece was used for convenience, containing two crossing lines at 25° with the central hair line.

All five cascades had lengths per plate of more than 1 cm. In the measurement on these cascades it is very difficult simultaneously to record all the dip angles : thus only tracks with a length per plate more than 3–4 mm were included. This corresponds to a dip angle of less than about 10°–15° with respect to the cascade.

Taking the centroid of all tracks in the target diagram as the cascade centre (i.e. the point where the core intersects the target diagram), the distribution N of particles inside the disc of radius R was obtained, after corrections for background were made. N was then plotted as a function of $Z_0 = E_0 \times R$. This result is presented in the graph. (For the con-venience of emulsion work the unit for Z_0 chosen was BEV $\times \mu$, i.e. R was measured in microns and E_0 in BEV).

The most remarkable feature about the graph is that at each value of t, ranging from 2 to 7 cascade units, the points for all five cascades fall on the same curves, although the primary energies vary sometimes by a factor of 10. This can be seen most clearly at $t = 4$ cascade units. Note the fluctuations at 3 cascade units. These are to be expected (Janossy 1950), since the fluctuations tend to increase towards either side of the maximum of the cascade. At the maximum they are believed to be nearly Poissonian. For the illustration of Janossy's expression ' fluctua-tion in effective depth ' the distribution of a cascade at 2·4 cascade units was drawn as a dashed line in the distribution after 3 cascade units. Note that the cascade with the full circles in the graph seems to show a fluctuation to ' smaller effective depth ' after 3 cascade units.

It may be concluded from the figure that the distribution near the core of the cascade is, apart from inherent fluctuations, a function of Z_0 only.

The full curves drawn in the figure are the result of a saddle point calculation of the theoretical distribution multiplied by a suitable factor. This factor is given in a circle at the end of each distribution. Therefore, although the shape of the theoretical function seems to be correct above 2 cascade units, there appears to be a discrepancy with regard to the normalization.

This discrepancy can have two causes : either the number of particles calculated from Nishimura and Kamata's theory of lateral spread is incorrect ; or the values of the primary energy E_0 deduced from the shower theory of longitudinal spread under ' Approximation A ' are in error. In the table details are given of two nuclear events, in each of which it was possible to associate pairs of photons with parent neutral pions, and in this way to obtain independent measurements of the primary energy E_0. The values of E_0 so deduced are seen to be in good agreement with those obtained from comparison of the lateral spread of the cascades with the graph. Thus, the values of E_0 calculated from the longitudinal
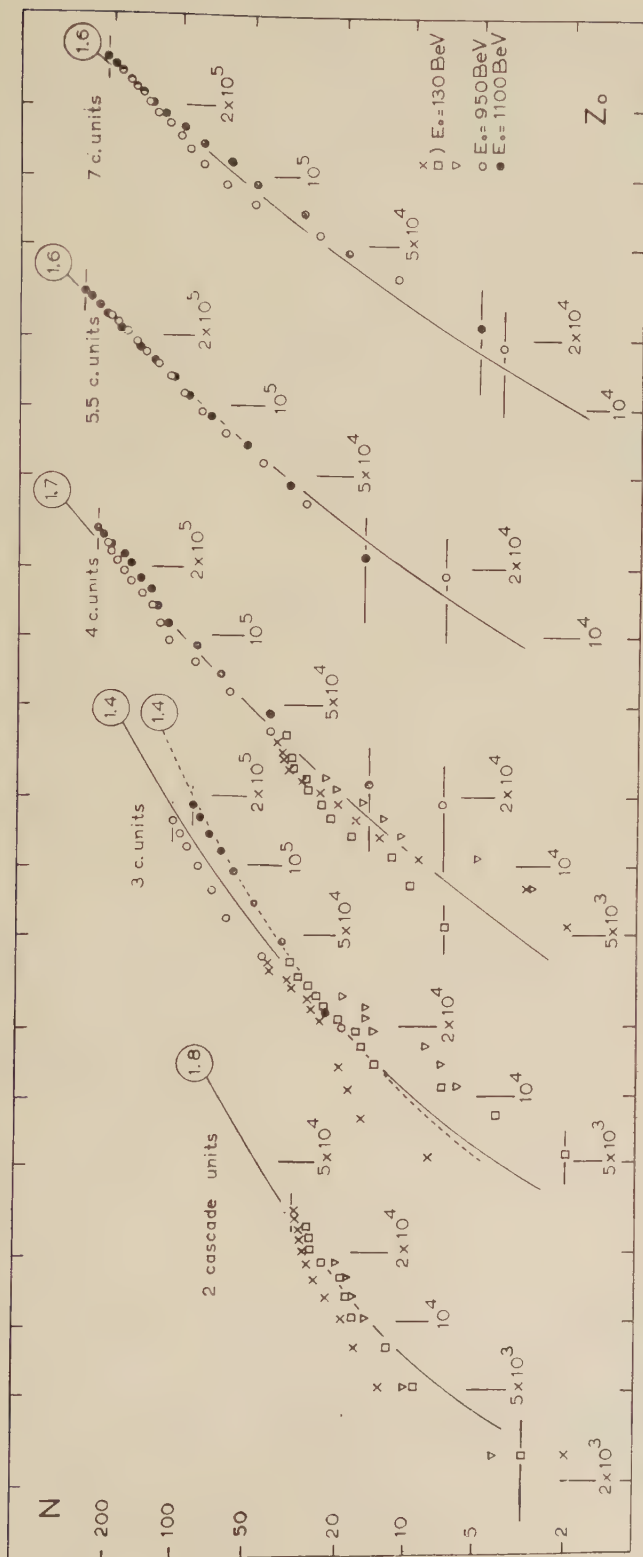
N, the number of tracks inside the disc with radius R is plotted as a function of the variable $Z_0 = E_0 \times R$, where E_0 is the primary energy of each cascade in BeV, and R is measured in microns. Only horizontal error lines are drawn in the graph. They represent the larger of the two error sources. (a) The uncertainty in the cascade centre, ——an error ΔR of $\pm 10\,\mu$ is given to the point with the lowest value of R in the distribution. (b) The uncertainty in E_0, ——an error of $\pm 20\%$ was assumed and drawn on the point with the highest value of R in the distribution. Statistical errors associated with the number of tracks, N, will be somewhat greater than Poissonian, and are not shown. The full lines represent theoretical calculations after Nishimura and Kamata's (1952) theory, the absolute numbers being multiplied by the factors shown inside the circle attached to each curve. The dashed line at $t = 2\cdot 4$ cascade units shows a distribution calculated for a value of $t = 2\cdot 4$ cascade units.

cascade development appear to be correct, and the error must therefore occur in the theoretical calculations of the lateral spread.

To summarize, the procedure proposed here for the energy determination of cascades is to measure the distribution of N *vs* R. Shifting this distribution along the Z_0 axis of the figure to obtain the best fit one relates a set of R values to a corresponding Z_0 scale. The ratio of Z_0 to R yields the primary energy E_0.

In conclusion the author would like to express his thanks to his colleagues of this laboratory, especially to Dr. D. H. Perkins, and to Mrs. N. Hillier for helping with the target diagrams. Dr. J. Iwadare kindly translated the letters of Dr. J. Nishimura, to whom thanks are due for his helpful explanations. The author is very much indebted to Professor C. F. Powell for the facilities of this laboratory, and to Professor E. Bagge and the Studienstiftung des deutschen Volkes for making this stay at Bristol University possible.

Jet	Number of pairs	Energy of the π^0's from opening angle of their γ-rays	Energy of the cascades as determined from the graph
0+2p	2	1100 bev	1200 (2000) bev†
0+4p	4	4100–4800 bev	5500 (9000) bev†

† The values in brackets refer to the energies E_0 when determined by using the theoretical curves without the correction factor.

In the case of the 0+2p star the measurement of the cascade suffered from the high background. Equipartition of the π^0-energy between the two γ-rays was assumed for the calculation of the value 1100 bev.

In the 0+4p star, the limits given for the energy of the π^0's enclose all possible combinations of the 4 pairs. The energy of this cascade was obtained after corrections had been made for a secondary jet occurring 7 cm before the point of measurement. It was assumed that this jet gives rise to about 30 particles within 200 μ distance from the core, and these particles were subtracted from the target diagram.

References

Brisbout, F. A., Dahanayake, C., Engler, A., Fowler, P. H., and Jones, P. B., 1956, *Nuovo Cim.*, **3**, 1400.

Janossy, L., and Messel, H., 1950, *Proc. phys. Soc. Lond.* A, **63**, 1101.

Nishimura, J., and Kamata, K., 1952, *Progr. theor. Phys.*, **7**, 185.

Pinkau, K., 1956, *Nuovo Cim.*, **3**, 1285.

Rossi, B., and Greisen, K., *Rev. mod. Phys.*, **13**, 240.

Hall Effect and Magneto-resistance in Indium–Antimonide

By Gaston Fischer and D. K. C. MacDonald
National Research Council, Ottawa, Canada

Recently published experiments (e.g. Chambers 1956, McDonald 1956, 1957) on the Hall effect and magneto-resistance in the Group I metals have emphasized that much still remains to be understood in this field (cf. also Ottawa 1956). The use of liquid helium and pure metal specimens enables one to achieve high values of the ratio l/r where l is the conduction-electron mean free path and r the orbital radius of a free electron in the applied magnetic field H. Olsen (Olsen and Rinderer 1954) has also used condenser discharge through a small coil to provide rather high magnetic fields of short duration.

The very high mobility of the negative carrier in InSb implies that even at room temperature one should have $l/r \sim 1$ for quite moderate magnetic fields. Using a large electromagnet with conically-tapered pole-pieces and about $\frac{1}{2}$ in. spacing we have been able to employ fields approaching 40 k Oersteds, which around $0^\circ c$ implies $l/r \sim 10$. We therefore decided to make some measurements on specimens of rather pure InSb prepared in our laboratories to determine whether the anomalous behaviour observed in metals when l/r becomes large is also to be found in a semiconductor.

Figures 1 (*a*) and 1 (*b*) show some typical results obtained in a first series of experiments for which no great accuracy is claimed. Wilson (1953) gives expressions for the Hall coefficient, A_H, and resistivity, ρ_H, of a two-band conductor in their dependence on H (cf. also Chambers 1952). If the densities of electrons and holes be n_1 and n_2 with corresponding mobilities μ_1 and μ_2 assumed independent of H, we can then show that

$$\frac{A_H - A_0}{A_0} \bigg/ \frac{\rho_H - \rho_0}{\rho_0} = \frac{(n_1 - n_2)\mu_1\mu_2}{n_1\mu_1^2 - n_2\mu_2^2}.$$

This ratio should thus be independent of H, although A_H and ρ_H may individually vary very strongly with H.

Figure 2 shows the results of plotting the experimental values of A_H and ρ_H in this way where some indication of possible limits of experimental error is shown. It is clear that the experimental ratio is not *strictly* constant and independent of H, but fig. 2 indicates that the two band model and theory based on the Boltzmann equation with $(df/dt)_{\text{collisions}} = -(f - f_0)/\tau$ offers a quite good first approximation for the behaviour of InSb even when $l/r \gg 1$.
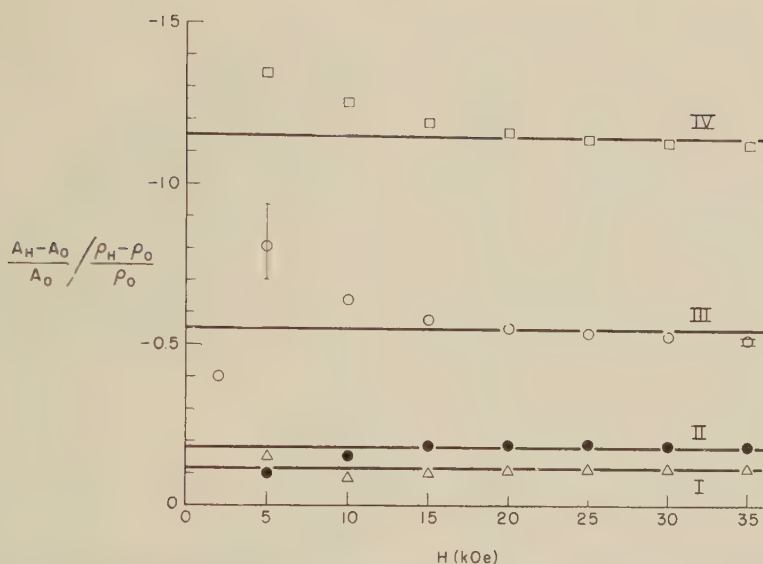
Fig. 1
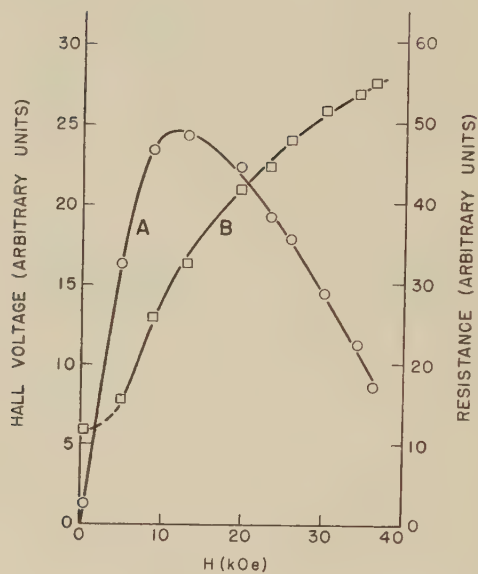
(a)
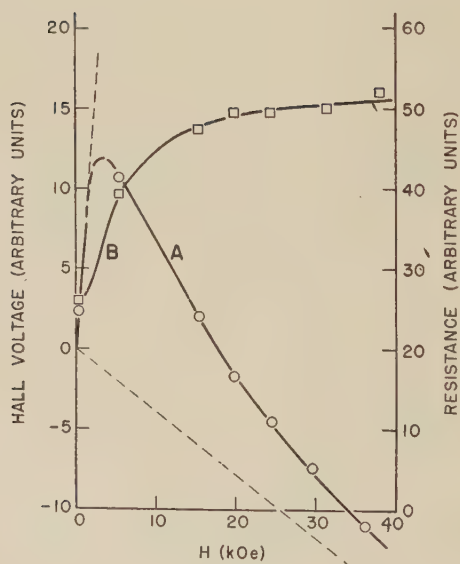
(b)

Measurements of Hall voltage (*A*) and electrical resistance (*B*) of InSb strip.
(a) at about 0°c ; (b) at about −12°c.

Fig. 2

Quotient of relative change in Hall constant and relative change in resistivity $\{[(A_H - A_0)/A_0]/[(\rho_H - \rho_0)/\rho_0]\}$ against magnetic field (H). I, at about 26°c; II, at about 0°c; III, at about —4°c; IV, at about —12°c. The horizontal bars applied to two points on curve III give an indication of possible limits of experimental error.

This work has been extended to cover a wider temperature range with more detailed measurements. We intend to make a more comprehensive analysis of the results and hope to publish more fully later in the *Canadian Journal of Physics*.

REFERENCES

CHAMBERS, R. G., 1952, *Proc. phys. Soc. Lond.* A, **65,** 903 ; 1956, *Proc. roy. Soc.* A, **238,** 344.

MacDONALD, D. K. C., 1956. Annexe 1956–2, *Supp. Bull. de l'Inst. Int. Froid*, p. 15 ; 1957, *Phil. Mag.*, **2,** 97.

OLSEN, J. L., and RINDERER, L., 1954, *Nature, Lond.*, **173,** 682.

OTTAWA, 1956, Proc. Int. Conf. on Electron Transport, *Canad. J. Phys.*, **34,** December 1956.

WILSON, A. H., 1953, *The Theory of Metals* (Cambridge : University Press), 2nd Edition.

REVIEWS OF BOOKS

Progress in Nuclear Physics. Vol. 5. Edited by O. R. FRISCH. (Pergamon Press, Ltd.) [Pp. 325.] 80*s.*

PROGRESS IN NUCLEAR PHYSICS is by now a well established member of the family of review journals, and the fifth volume in the series maintains the high standards set by its predecessors.

The emphasis in the present volume is on experiment. Four out of the seven articles are concerned exclusively with the tools and methods of the experimental nuclear physicist. One describes deflection methods for the accurate determination of nuclear reaction energies. Another gives an account of some of the electronic techniques that have been developed in the last ten years, while the other two are concerned with the bubble chamber and organic scintillators.

There is an account of the knowledge that has been obtained from the inelastic scattering of fast neutrons and also an article which discusses the concept of the size of the nucleus and analyses the results of the various methods of measuring this quantity. Finally there is an article on the neutrino. Written before the recent discovery of the non-conservation of parity in weak inter- actions, it gives the experimental evidence for the neutrino and the theoretical position as it appeared at about the beginning of 1956.

All the articles are well written, have clear diagrams and contain detailed bibliographies. The only adverse criticism this reviewer has of the book is that its price must surely be beyond the pockets of all but a handful of students and research workers. G. L. S.

Progress in Semiconductors. Vol. 2. Edited by A. F. GIBSON, P. AIGRAIN and R. E. BURGESS. (London : Heywood & Co.) [Pp. 280.] 63*s.*

VOLUME 2 of this series follows the pattern set in the first volume. There are eight essays on topical aspects of semiconductor physics, five of which are either solely concerned with the special properties of germanium, or lean heavily on them for illustrative purposes. The editorial policy of combining review articles with original research is not uniformly successful because this tends to over-emphasize some parts of the subject matter. The result is that while the articles contain a wealth of up-to-date detail they do not always carry a strong sense of conviction.

There are some interesting speculations about the band structure of disordered crystals in Herman's article on semiconductor alloys, and a courageous attempt by Rose to bring order into photoconductors containing several different types of recombination centre. More conventional are the reviews of III–V compounds by Cunnell and Saker, and impurities in germanium by Dunlap, both of which are competently treated. O. S.

ERRATUM

The Effect of Free Electrons on Lattice Conduction at High Temperatures, by R. STRATTON, 1957, *Phil. Mag.,* **2,** 422.

In eqn. (5) replace kT by $\hbar\omega$ in the numerator.

Fig. 6   Fig. 7

0·25μ   ×50 000   0·25μ   ×60 000

Fig. 8

0·25μ   ×60 000

Fig. 9   Fig. 10

0·25μ   ×50 000   0 25μ   ×50 000

Fig. 11

0·25μ   ×50 000

Fig. 13

(a)

(b)

(c)

(d)

0·25μ

×40 000

Fig. 14

0·5μ   ×30 000

Fig. 15

(a)   (b)

(c)

Fig. 16

Fig. 17

0·5μ   × 30 000   1μ   × 10 000

Fig. 18

Fig. 19

0·5μ   × 20 000

Fig. 22

0·25μ   (a)   (b)   ×50
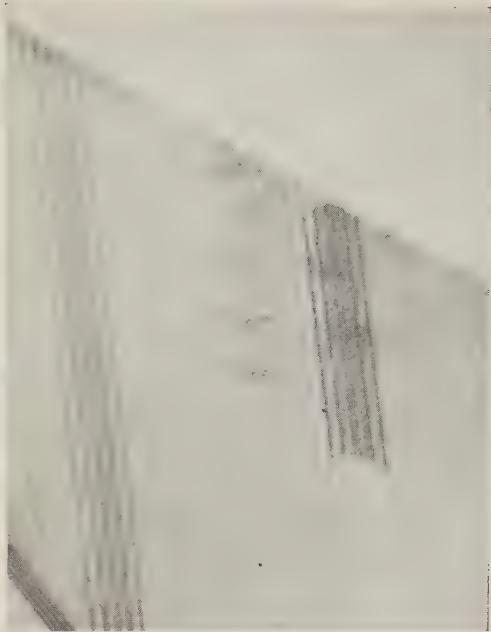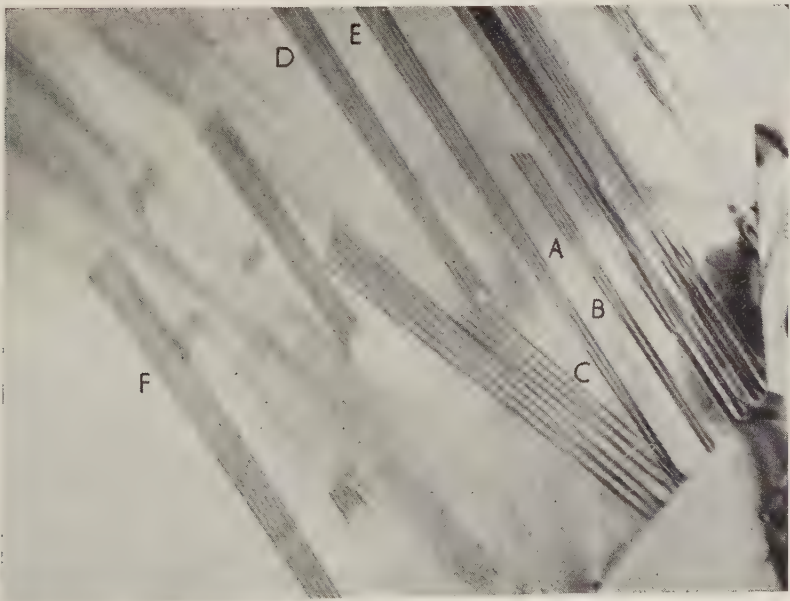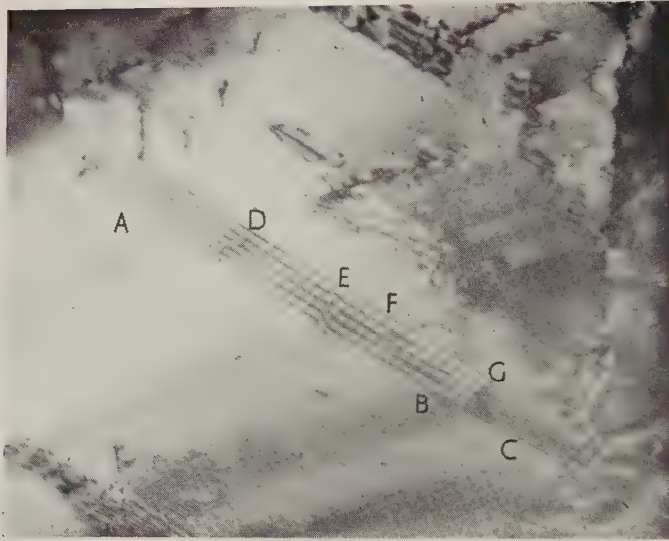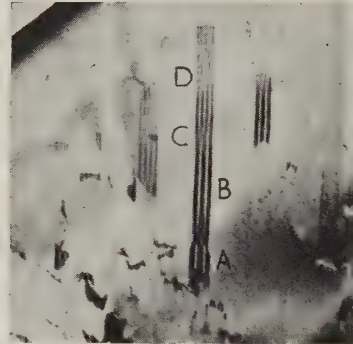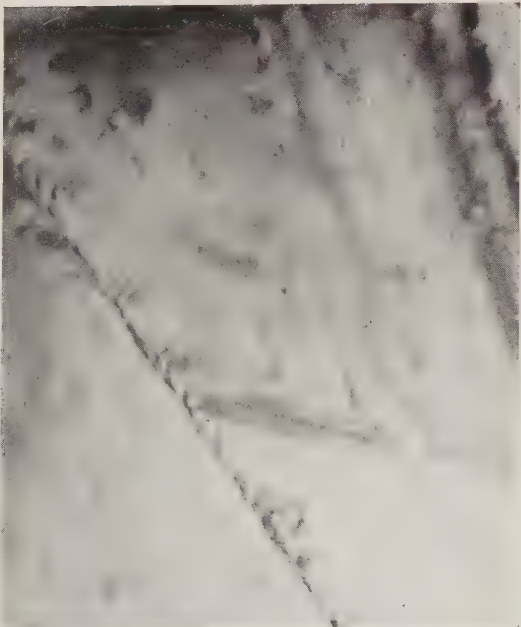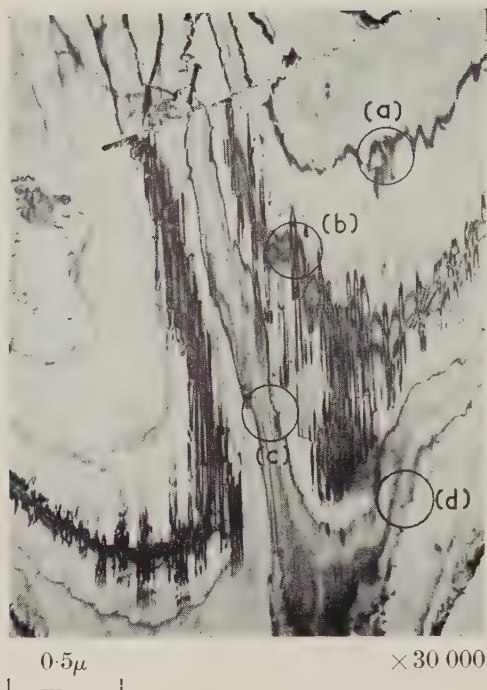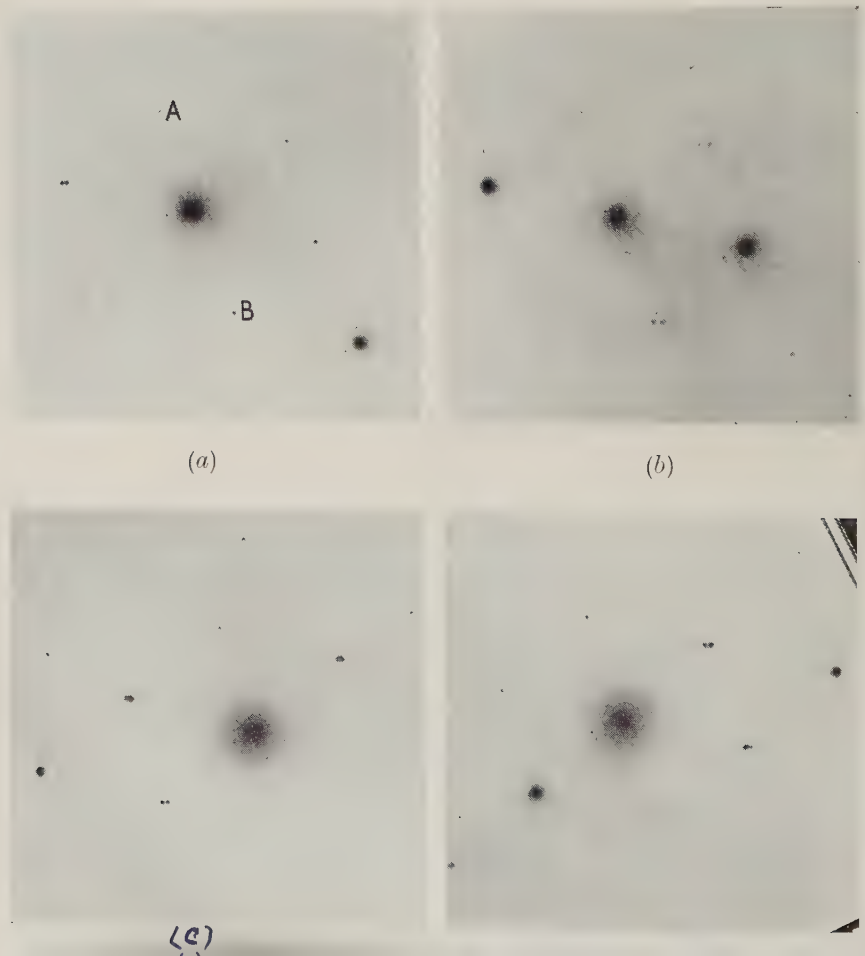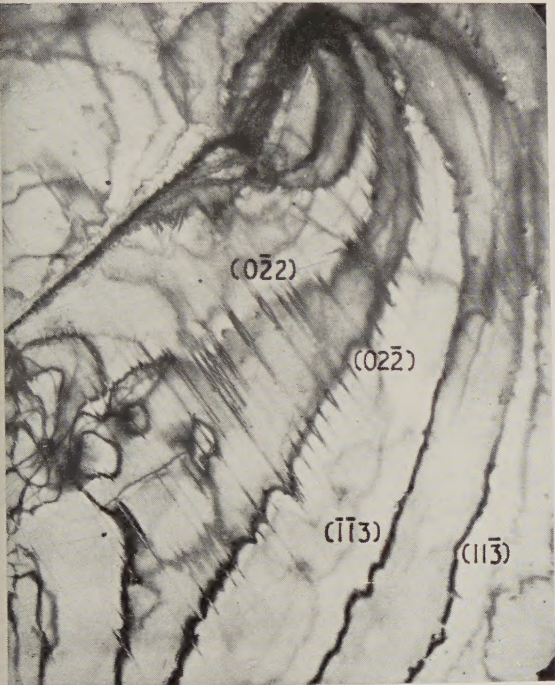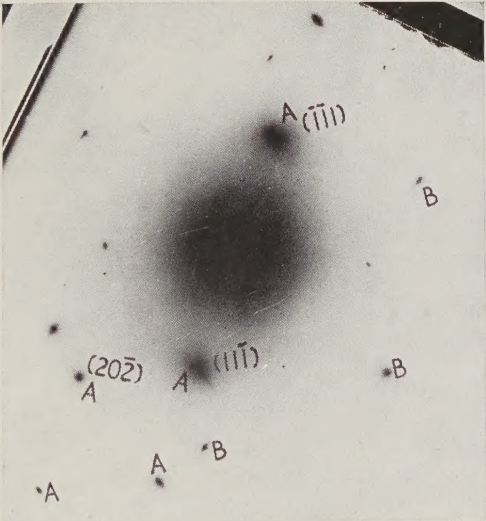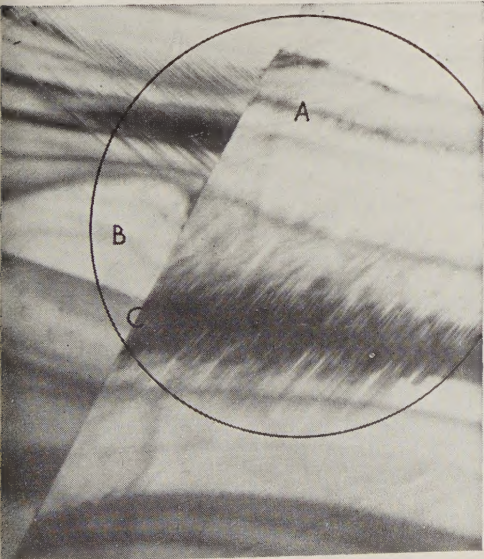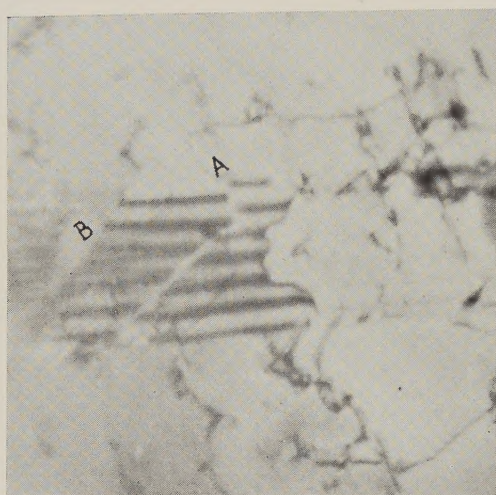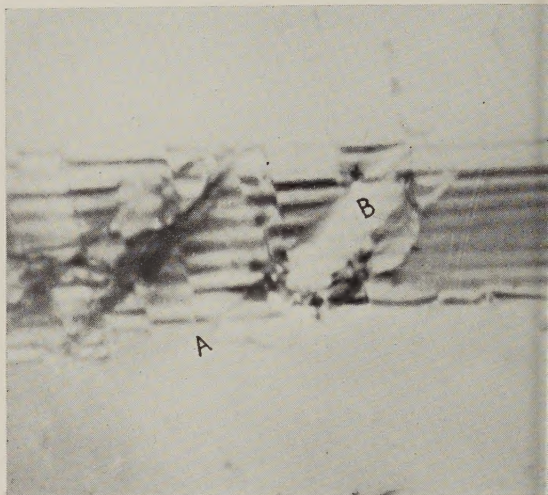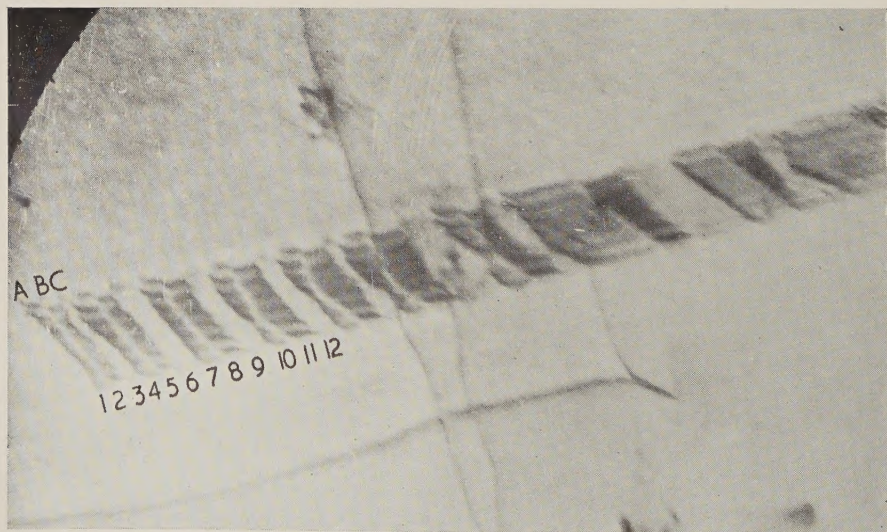
Fig. 23

0·25μ   ×40 000

CAPTIONS TO PLATES

Fig. 6

Fringes from two stacking faults side by side. These fringes are character-istic of the dynamical region ; note the doublet nature of the dark fringes. Magnification ×50 000.

Fig. 7

Fringes typical of those due to faults produced during observation in the micro-scope. Note the absence of the doublet structure at the edges of the faults. Magnification ×60 000.

Fig. 8

Variation of fringe profile with thickness. The varying thickness of foil in this region is apparent from the taper of the faults, e.g. at D, E, F. Weak subsidiary maxima of extinction contours can be seen at A, B, C. Close to where these meet the faults, reversal of the fringes occurs. Mag-nification ×60 000.

Fig. 9

Effect due to changes in orientation. A corresponds to the kinematical region ; at B the doublet structure appears in the dynamical region ; at C, $x \simeq 0$. Dislocations on neighbouring planes are responsible for the features D, E, F. Magnification ×50 000.

Fig. 10

Variation of number of fringes with x. At A, x is small ; the doublet structure is apparent. At B, x is larger and the fringes are single. Reversals take place at C and D, the number of fringes increasing by one at each reversal. Magnification ×50 000.

Fig. 11

Fringes typical of thin regions ($t \sim t_0$). Near the strong extinction contours the fringes are in the form of one or two bands. Magnification ×50 000.

Fig. 13

Sequence of micrographs showing the effect of tilting the specimen on the fringe contrast. The axis of tilt and the relative positions of the (111) and ($\overline{1}\overline{1}\overline{1}$) reflections are indicated in (a). See text for detailed discussion. Magni-fication ×40 000.

Fig. 14

Area showing many stacking faults and extinction contours. The circles (diameter 0·25 μ) mark the approximate areas selected for the diffraction patterns of fig. 15. Magnification ×30 000.

Fig. 15

Diffraction patterns from the corresponding areas indicated in fig. 14, straddling different extinction contours. These patterns are used to index the contours. In (a) doublet spots corresponding to weak (111) reflections can be seen clearly at A, B.

Fig. 16

Micrograph showing fringe contrast at contours $02\overline{2}$, $0\overline{2}2$, but not at $\overline{1}13$, $11\overline{3}$. Magnification ×30 000.

Fig. 17

Low magnification picture of same area as that of fig. 16, showing that faults occur in the whole region between the ($\bar{1}$11), (1$\bar{1}$$\bar{1}$) and (1$\bar{1}$$\bar{1}$) ($\bar{1}$11) contours. Note the relatively greater contrast of the fringes near the (111) type contours compared to that in the (022) type contours. Magnification $\times$ 10 000.

Fig. 18

Area showing large numbers of stacking faults in two grains A, B separated by the grain boundary C. The circle represents the selected area from which the diffraction pattern of fig. 19 is obtained. Magnification $\times$ 20 000.

Fig. 19

Electron diffraction pattern from the area indicated on fig. 18. Spots labelled A and B come from the grains labelled A, B, respectively on fig. 18. The streaks in this pattern run parallel to directions consistent with stacking faults on (111).

Fig. 22

Overlapping stacking faults. At A the predicted fringe reversals can be seen. At B no contrast is observed ; This may be due to overlap of three faults. Magnification $\times$ 50 000.

Fig. 23

Effect of overlap of a number of stacking faults. The fringes at A are due to a single fault. At B the fringes reverse owing to overlap of two faults ; at C the contrast disappears probably owing to overlap of three faults. This sequence repeats and is probably due to a set of overlapping faults ; the number of overlapping faults is indicated on the figure. Magnification $\times$ 40 000.